

THE MODELING OF DAILY PRECIPITATION IN COSTA RICA

BY

JOHN MICHAEL HARRISON

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1998

For my father, the late Marshall Gene Harrison, who pointed me on the path I could never have taken otherwise; and my mother, Mary Ann Harrison, who made sure that I never strayed from it. My thanks and love, always.

ACKNOWLEDGEMENTS

A dissertation can never be the product of a single individual, and I regret that it is not possible to list as co-authors all the people that have helped make this work possible. I am grateful, on both a personal and professional level, for the support and friendship offered by Mark McLean, Chris and Kathleen Meindl, Matt and Mary Zorn, Sharon Cobb, Ryan Jensen, and Lisa Walsh. These fine individuals have made life as a graduate student not simply bearable, but enjoyable, and I am proud to call them friends.

I have been blessed with an outstanding advisory committee, and without their advice and support I could not have survived the arduous process of completing my degree. I am extraordinarily grateful for the support lent by Dr. Joann Mossa, Dr. Michael Binford, and Dr. Jim Jones for their expert guidance throughout my career as a doctoral student. I am especially grateful to Dr. César Caviedes, who not only illuminated the mysterious ways of El Niño, but also allowed me the opportunity to teach within the Department.

Most of all, my thanks go out to Dr. Peter Waylen, who has proven to be the finest advisor I have ever encountered. It is safe to say that without his help, insight, and encouragement, I could not have completed my degree. Throughout this journey, he has been there to help me, in good times and bad, and I can only hope that this work, and all those that follow, will in some small way repay the confidence he has shown in me.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS	iii
ABSTRACT	vi
CHAPTERS	
1 INTRODUCTION	1
Rainfall Modeling	2
General Characteristics of Costa Rica	4
Structure of this Dissertation	5
2 LITERATURE REVIEW	8
Probability Models	8
Forecasting and Prediction	15
El Niño-Southern Oscillation	17
3 STUDY AREA AND DATA DESCRIPTION	24
Physiography	24
General Climatic Characteristics	25
Previously Noted Costa Rica-ENSO Connections	26
Daily Precipitation Data and Station Location	29
Atmospheric Index Data	30
4 METHODOLOGY	47
Precipitation Statistics	48
Relationship to ENSO Processes — Segregated Analysis	56
Relationship to Ocean-Atmosphere Processes — Lag-Correlation Analysis	56
Creation of Statistical Forecast models of Rainfall Parameters	57

5	RESULTS.....	61
	Model Identification.....	61
	Time Series Correlation of Transition Probabilities.....	67
	Time Series Correlation of L-Moments	72
	Model Choice, Validation, and Rainfall Simulation	76
	Results Overview	91
6	DISCUSSION	182
	Model Identification.....	182
	Time Series Correlation — L-Moments and Transition Probabilities	194
	Model Choice, Validation, and Rainfall Simulation	202
	Parameter Forecasting	205
7	CONCLUSION	212
	Statistical Modeling.....	213
	Climatology	215
	Application to Other Areas.....	220
	APPENDIX L-MOMENTS.....	222
	REFERENCES	228
	BIOGRAPHICAL SKETCH	240

Abstract of Dissertation Presented to the Graduate School of the University of Florida in
Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

THE MODELING OF DAILY PRECIPITATION IN COSTA RICA

By

John Michael Harrison

August 1998

Chairman: Dr. Peter R. Waylen
Major Department: Geography

The understanding of precipitation and its underlying processes is important to many human activities. Agricultural planning, hydroelectric resource management, and industrial infrastructure development all rely heavily on being able to make reasonable predictions concerning rainfall. The lack of sufficient rainfall can have devastating social and economic consequences for developing nations that are reliant on subsistence agriculture and hydroelectric power. This study examines the means by which daily precipitation in Costa Rica can be modeled, and how the El Niño-Southern Oscillation (ENSO) affects the precipitation generating mechanisms.

A selection of three meteorological stations are used to test how daily rainfall can be characterized by the occurrence and intensity of the individual rainfall events. The occurrence is modeled using a two-state first-order Markov model, which provides insight into the relative length of wet and dry spells. The intensity model uses L-moments to determine the optimum statistical distribution. These statistical parameters are used to

understand the inter-annual and inter-seasonal variations in the precipitation-generating mechanisms as they are modified by the ENSO phenomenon. The parameters are also combined to create monthly rainfall simulations based on the state of the ENSO, as well as test whether accurate forecasts can be created up to one year in advance.

It is found that the ENSO plays an important role in the daily rainfall process, by altering the behavior of the Inter-Tropical Convergence Zone (ITCZ), the Northeast Trade Winds, and the advance of cold air masses from North America during the winter. The eastern Caribbean slope of the country receives proportionally more rainfall during El Niño events, while the western Pacific slope receives less rainfall during the same period. Cold front (*norte*) intrusion is minimized by the El Niño, resulting in less winter rainfall during El Niño years. Simulations based on the state of the ENSO are shown to be effective in recreating monthly time series of rainfall accumulations based upon the calculated occurrence and intensity parameters. Long-range forecasts are less effective, which is indicative of the complexity of the tropical precipitation process.

CHAPTER 1 INTRODUCTION

One of the most important of all climatic variables is precipitation. It touches on nearly every aspect of human existence, both from a social and economic standpoint and as a field of scientific study. Human societies rely heavily on precipitation as a necessary input to agricultural systems and the resultant runoff has for centuries driven machines, and more recently generated electricity. At the same time, water has had destructive side-effects, such as floods, soil erosion, mass movement, and the disruption of transportation and communication infrastructures. Because humans are so reliant on rainfall, the ability to predict accurately its occurrence and intensity has been pursued with exceptional energy. In many parts of the developing world, the understanding of rainfall occurrence and intensity goes well beyond mere convenience. Much of the world's population continues to survive on subsistence farming systems, and many developing nations depend heavily on the export of cash crops and the generation of hydroelectric power. For these reasons, fluctuations in the amount of rainfall in these, such as those prompted by the El Niño-Southern Oscillation phenomenon, can have dramatic and devastating economic and social results.

The nation of Costa Rica is an excellent case-in-point. It is a small Central American country attempting to deal with the often conflicting needs of a highly agrarian economy moving toward a more industrialized status with a burgeoning population that is becoming increasingly urban. Over 60% of national electrical power is derived from hydroelectric sources, and is therefore highly sensitive to changes in the precipitation climatology of the region. In addition, the agricultural systems that represent 60% of the

GNP are heavily reliant on rainfall, with less than 25% of all cropland subject to irrigation. Unlike many countries in the region, however, Costa Rica has a stable economic and social system, and this has allowed it to develop an extensive network of meteorological stations, with relatively long-term records.

This research will use a set of three rainfall stations in the Central Valley of Costa Rica to test the effectiveness of daily rainfall modeling and prediction within a tropical context. Specifically, it will be shown how the El Niño-Southern Oscillation (ENSO) affects the intensity and frequency of daily rainfall over the study area, and how the daily rainfall characteristics can be modeled with a high level of confidence. The creation of such models will represent a substantial benefit to climate researchers, agricultural engineers, and planners.

Rainfall Modeling

A wide variety of approaches exists in the study of precipitation and precipitation-generating processes. One, usually associated with meteorologists, examines causal physical processes, and attempts to extend the findings in such a way as to explain spatial and temporal variations in precipitation. While such an approach is useful in illuminating the physical processes involved, it can lead to conclusions that are spatially and temporally limited. Physical processes are usually complex, and the resulting models are either unacceptably complex (i.e., lacking in parsimony, with many variables involved, thus making it difficult to discern the most important processes involved), or overly restrictive, thus preventing generalization of their results. Intensive data demands also make them generally less applicable to most regions of the world.

An alternate approach is to examine the time series of the end-result (i.e., the precipitation itself), as a means to understanding its pertinent properties. The rainfall

record then becomes the integrated output of all the contributing physical processes. This is an especially useful approach when the exact physical mechanisms are not well understood, are complex, and/or pertinent data are difficult to obtain. The statistical approach may mitigate many of the problems associated with modeling causes of climatic variability such as the ENSO, while maintaining the goal of making accurate predictions. In general, most statistical analyses of daily rainfall seek to model the occurrence of precipitation (as a sequence of wet or dry days), and the probability distribution of the intensity of the resulting precipitation.

From the standpoint of statistical modeling, this study seeks to address three broad issues. First, the nature of the seasonal and spatial variations of daily rainfall over the Central Valley of Costa Rica will be examined. Models of occurrence and intensity of precipitation are to be developed, using rainfall records available from the appropriate agencies in Costa Rica. Second, the effect of the ENSO phenomenon on the spatial and temporal variability of the region's rainfall will be examined through its influence upon model parameters. Previous studies (e.g., Estoque et al., 1985; Fernández and Ramírez, 1991; Waylen et al., 1994) have demonstrated a link between Costa Rican annual and monthly rainfall and the state of the ENSO as represented by the so-called Southern Oscillation Index (SOI). It is important to establish the spatial extent and nature of the associations at the daily scale. Third, the ability of these models to predict accurately future daily rainfall characteristics will be examined. Several modeling techniques will be employed to create "hindcast" predictions, and the results will be compared to the measured rainfall during these periods.

This study yields statistical models with broad applicability to a variety of regions in Costa Rica and Central America that will help to illuminate the nature of the physical processes which affect precipitation in the tropics, particularly the ENSO phenomenon. Ultimately, correlations between daily rainfall characteristics and the indices describing the ENSO, such as SOI, could lead to a method for predicting future rainfall variation based upon the state of the ENSO. This will have broad applications for both short-term precipitation forecasts and long-range climate prediction.

General Characteristics of Costa Rica

The nation of Costa Rica occupies an area of 51110 km², supporting a population of 3,032,000 (Goode's World Atlas, 1992) (Figure 1-1). Its economy is strongly linked to agriculture, but the trend has been towards increased urbanization, with well over half the population now living in urban areas (West and Augelli, 1989). This has placed increasing demands on the extensive hydroelectric projects that are being developed (hydroelectric power accounts for about 60% of the total generated electricity in Costa Rica), as well as pressing for increased availability of water for drinking and industrial concerns (Gleick, 1993). In addition, despite a shrinking rural population, rural access to sanitation services has fallen from 93% in 1975, to 89% in 1985, reflecting the increased competition between rural and urban populations for available water resources. The Costa Rican agricultural system is strongly dependent on regular, predictable rainfall, with rainfed cropland accounting for over 77% of cultivated land (Gleick, 1993).

Costa Rica has had a relatively peaceful history and stable government, and this social environment has allowed for the creation of a reasonably diverse and well-supported meteorological network. Reliable precipitation records in this region date back

more than 150 years in some cases. It has also been shown that, due to its geographic position in the tropics and its susceptibility to meteorological influences from the mid-latitudes, Costa Rica experiences considerable spatial variability in precipitation in response to the state of the ENSO (Waylen et al., 1996).

Structure of this Dissertation

This research is organized along the following lines. First, previous research into the nature of the statistical modeling of rainfall magnitudes and intensities will be recounted. This will be followed by a brief history of forecast modeling, with particular concentration on the ARIMA and multiple regression techniques currently in widespread use. Next, the evolution of the current understanding of the El Niño-Southern Oscillation (ENSO) will be presented, with particular emphasis on its impact upon Central American climatology. In addition, a discussion of the atmospheric indices to be used in this study will be presented.

The study area is delineated in considerable detail in Chapter 3, especially emphasizing the known relationships between the spatial variability of Costa Rican rainfall and the ENSO phenomenon. The major climatological elements are introduced and discussed, as well as the measurement indices by which they are represented. The individual meteorological stations used in the study are evaluated for applicability, both in terms of geographic location and record length. Specific stations are identified for detailed examination, based upon their suitability to represent broad regions within the country.

A research methodology is then presented, which will outline the creation of daily rainfall intensity and occurrence models to be used in this study. The basic occurrence parameters are calculated using a Markov model, while the intensity statistics are computed using linear moments, or “L-moments.” The resulting model parameters are examined in terms of whether the statistic occurred during an ENSO event, with the intention of evaluating the broad-scale influence of ENSO upon Costa Rican daily rainfall. Next, lag-correlations between several atmospheric indices and the calculated statistics will be performed, so that more subtle patterns of influence can be detected. Finally, the parameters will be used as the basis for a series of statistical hindcast forecast models, which will be evaluated on the basis of how well they approximate known rainfall conditions for periods of up to one year into the future.

Finally, the results of the study will be evaluated in terms of the statistical importance of the correlations and forecasts at the individual stations, as well as in the broader context of regional climate variability. The model results will be placed into a format that can be directly accessible by researchers in other areas, such as reservoir and crop modeling, and an attempt will be made to reconcile the statistical models for Costa Rica with the known factors affecting the climate of the Central American region. It is anticipated that this study will provide much needed information for researchers in many other fields, as well as shed light upon the underlying physical processes governing the precipitation process in Costa Rica.

Figure 1-1. Costa Rica political units
(source: Perry-Castanada Map Collection, University of Texas at Austin)

CHAPTER 2

LITERATURE REVIEW

Pertinent research relating to the understanding and modeling of daily precipitation and ENSO-based climate variability are discussed. The nature of chain-dependent precipitation models is first reviewed, followed by a discussion of continuous probability models which have been employed to model precipitation magnitudes. A brief history of the climatological impacts of ENSO is recounted, with special emphasis on its effects with Central American climate.

Probability Models

The modeling of precipitation has often been viewed as an almost intractable problem, as it is unknown whether a physically realizable upper limit to the amount of rainfall that can fall during a given event exists, for a given set of environmental parameters (Knisel et al., 1979; Willeke, 1980; Hirschboeck, 1988). In addition, it is impossible to fully quantify all the different possible physical states that can exist during a given precipitation event. For these reasons, deterministic models are generally regarded as inadequate when evaluating the precipitation climatology of a region (Dutton, 1986).

Modeling of daily precipitation climatology therefore reduces to deriving the proper distributional models for the occurrence and intensity of the precipitation (Coe and Stern, 1982; Woolhiser, 1992; Cong and Li, 1993). The occurrence is frequently represented using Markov models (Chin, 1977; Katz, 1977, 1981; Coe and Stern, 1982; Stern and Coe, 1983; Guzman and Torrez, 1985; Hughes and Guttorp, 1994; Lall et al.,

1996), while the magnitude is most frequently denoted by a variety of continuous probability models, such as the gamma, generalized extreme value (GEV), and Pareto distributions (Markovic, 1965; Haan, 1977; Hosking, 1985; Guttman, 1993; Hosking and Wallace, 1993).

Probability Models -- Occurrence

Access to meteorological data which has been discretized into regular intervals allows the variable to be examined in terms of its dependence upon its previous values. This implies that the data sequence is stochastic in nature so long as the series is not excessively burdened by long-term serial dependence (Wilks, 1995). Stated simply, each value of the data stream (e.g., daily precipitation totals) must exhibit reasonable independence from values observed long in the past. If this assumption is met, the data are said to have "Markovian properties," the process which generated the data is referred to as a Markov process, and the resulting data model being called a Markov chain model (Haan, 1977). An important characteristic of a Markov model is that the value of the current data point depends only upon a fixed number of preceding measurements, and is independent of any other data. This implies that, in order to construct a Markov model of rainfall, the data must be analyzed both in terms of its temporal characteristics (i.e., how often precipitation occurs), and in terms of the measurement of the quantity (e.g., rainfall amounts). The degree of temporal dependence is determined by the order of the model (a first order model contains dependence only upon the value immediately preceding it in time, and is assumed independent of all other values), while the number of possible categories or classes within the basic time unit is referred to as the state of the model. For example, it is common to categorize rainfall occurrence as either "rained" or "did not rain"; this would represent a two-state model of rainfall (Haan, 1977; Wilks, 1995).

The model is typically produced by calculating the *transition probabilities* from the data sequence, by calculating the number of wet and dry days and conditioning these results on the basis of the prior day's state. Thus, a two-state first-order Markov model would contain four parameters: p_{00} (the probability of a dry day following a dry day), p_{01} (the probability of a wet day following a dry day), p_{10} (the probability of a dry day following a wet day), and p_{11} (the probability of a wet day following a wet day). It is important to note that p_{00} and p_{01} will always add to one, as will p_{10} and p_{11} . The creation of a matrix of four transition probabilities therefore requires the computation of only two values, which improves the parsimony of the resulting model (Wilks, 1995).

Due to the fact that rainfall is often caused by conditions which linger for multiple days (or longer), a certain amount of serial correlation will be apparent in a Markov precipitation chain. This is known as persistence, and can be measured using the lag-1 autocorrelation parameter, also known as the *persistence* parameter r_1 , which is the difference between the p_{11} and p_{01} parameters, and reflects the degree of serial dependence of the model (Wilks, 1995). Perfect independence between days of rainfall would imply $r_1 = 0$, which generally does not occur in natural systems. This would seem to indicate that the necessary conditions for applying Markov models to rainfall occurrence have been violated, since having serial correlation from one day to the next would indicate a degree of correlation for days separated by k -days, according to the formula $r_k = (r_1)^k$. As Wilks (1995) points out, this does not violate the basic Markovian assumptions, which state that successive conditional probabilities be independent, rather than the actual data points.

Past research has indicated that, with few exceptions, Markov models of the first order (i.e., those, in which a day's precipitation occurrence is conditioned only upon the previous day's occurrence, and ignoring events prior to that time) are the most applicable models for rainfall modeling (Caskey, 1963; Katz, 1977, 1981; Coe and Stern, 1982); however, these studies involve the use of mid-latitude rainfall records, and little published research indicates whether this holds for tropical precipitation. Wilks (1995) demonstrates how maximum likelihood estimates can be created for the applicability of first order versus higher order Markov models to a given data sequence.

Markov models have previously been used extensively to perform realistic simulations of precipitation occurrence which are subsequently input into deterministic models, for purposes such as climate change impact assessments (Pickering et al., 1994; Wilby, 1994), as well as to gain insight into the underlying physical processes which govern rainfall occurrence (Katz, 1977). Refinement of these procedures have shown how periodic and quasi-periodic physical processes responsible for seasonality in precipitation can be assessed (Woolhiser and Roldan, 1986; Woolhiser, 1992; Woolhiser et al., 1993; Katz and Parlange, 1993; Hughes and Guttorp, 1994).

Probability Models -- Magnitudes

Deriving the underlying distributional form for the intensity (i.e., the daily totals) of precipitation has been more problematic, due to the open-ended nature of extreme precipitation events and the resulting highly skewed rainfall histograms (Woolhiser, 1992). A further complicating factor is the large number of candidate continuous distributions from which to choose. Many two- and three-parameter probability distributions are applicable to a variety of hydrologic and climatological problems;

among those that have found broad appeal are the exponential, Gumbel (Haan, 1977), log-normal (Cong et al., 1993; Stedinger, 1980), generalized extreme value (Haan, 1977; Hosking, 1985), gamma and Pearson type III (Cong et al., 1993; Haan 1977), Pareto (Cong et al., 1993; Hosking and Wallis, 1987), Weibull (Haan, 1977), mixed exponential (Woolhiser et al., 1993), and Wakeby distributions (Landwehr et al., 1979).

Most intensity models involve the manipulation of an exponential function; the rationale for this is that the exponential function can easily be made to represent a rapidly decreasing probability density function, while still preserving many of the desirable mathematical properties that make the exponential function so attractive (Haan, 1977). Examples of this type of function are the exponential, gamma, and Pearson distributions.

One of the major difficulties in working with precipitation data is the presence of a lower boundary on the data values; there can never be a negative precipitation, and zero amounts are modeled as "occurrence/non-occurrence" within the context of the Markov model. For this reason, distributions that are continuous in both the negative and positive directions are generally not applicable. However, it is possible, in some cases, to mimic the behavior of a normally distributed variable by taking the logarithm of the amounts, and using the transformed values as a normally distributed variable; this can often give acceptable results, and the resulting distribution is called a "log normal" (or "generalized normal") distribution (Haan, 1977). This procedure has been used in modeling hydrologic data such as stream discharges (Stedinger, 1980), but has found limited utility when modeling daily rainfalls due to the availability of other functional forms.

The lack of an effective upper limit on precipitation values forces the consideration of more complex distributions which can take into account highly skewed data. Most of these distributions are some variation on the exponential function. A pure exponential, however, is unusable because of its fixed skew (equal to $2/\text{mean}$); for a function to be useful, its skew must be capable of fitting a broad spectrum of data. A

more generalized version of the exponential is the “gamma function,” which has a variable skew; the basic gamma function has two parameters, which together define the variability and skew of the resulting probability distribution (Haan, 1977). Gamma distributions have been used extensively to represent rainfall amounts (Markovic, 1965; Mooley and Crutcher, 1968).

In hydrologic research in general, and more specifically, studies of precipitation, the extremes of a set of events is often of interest. The extreme value of a set of random variables is itself a random variable, with the resulting distribution depending on the sample size and the underlying distribution of the individual events; these are known as “extreme value distributions” (Haan, 1977). The most common type is the generalized extreme value distribution, which can be shown to be the asymptotic limiting distribution for a variety of exponentially based distributions with one tail unbounded. The lognormal, exponential, and gamma distributions all share this trait (Haan, 1977). The generalized extreme value distribution has three-parameters, allowing the location, scale, and shape to vary (Jenkinson, 1954; Hosking, 1985).

A very useful class of distributions was developed by Pearson (Elderton, 1953) which can be manipulated to form several well known variants. In particular, the Pearson Type III distribution represents a three-parameter generalization of the gamma distribution, and has found broad applicability to flood and precipitation peak research (Haan, 1977). By increasing the number of associated parameters, the distribution can better model the mode of a data set, allowing a better fit than if the mode were fixed, as in a two-parameter gamma.

Distributions can also be mixed, if there is reason to believe that the precipitation originates from more than one process. Often, these models can better represent the data, but they do so at the expense of parsimony; mixed models, those composed of two or more lower order distributions, have as many parameters as all of the underlying

component distributions. Conceivably, any set of distributions can be mixed to assemble new distributions, but understanding the nature and statistical properties of the resulting model can be difficult at best, and will always result in a larger number of parameters, thus reducing the total degrees of freedom of the calculations. Examples of such models are the mixed exponential distribution (Woolhiser et al., 1993); and Wakeby distribution, which is simply a mixed Pareto model (Landwehr et al., 1979).

Parameter estimation. The traditional means of estimating the underlying distribution parameters has been to employ the “method-of-moments” approach, in which the statistical moments (mean, standard deviation, etc.) are computed from the data record (Wallis et al., 1974; Stedinger, 1980). However, this method is handicapped by the fact that the extreme precipitation values tend to be poorly represented (Vogel and Fennessey, 1993; Harrison, 1995). Since these extremes are of critical importance in many practical applications, more efficient parameter estimation methods are needed.

Recently the use of “linear moments,” or L-moments, originally developed to deal with the problem of highly skewed data sets (Greenwood et al., 1979), have shown some promise in the realm of climatological studies. They have the advantage of producing distribution parameters that are more stable than those derived with the traditional method-of-moments approach (Hosking, 1989, 1990; Hosking and Wallis, 1993; Harrison, 1995). The reason for this is that, as the name implies, they are linear combinations of probability-weighted moments, rather than power-driven combinations, as with traditional moments. In calculating the values of traditional moments such as variance, skew, and kurtosis, the variation from the mean is raised to a power, thereby causing outlying values to exert an inordinate amount of influence, leading to unstable results with highly skewed data such as daily precipitation. It was shown by Hosking and Wallis (1993) and Harrison (1995) that L-moments can give more stable results than traditional moments when rainfall data are used. Mason et al. (1996a, 1996b) demonstrate

that L-moments highly effective when used to estimate the return periods for rainfall and changes in flood frequencies over South Africa. Guttman (1994) further illustrates that L-moments can be used with confidence for the determination of skew and kurtosis if as few as 60-70 samples are available, making the method well suited for long-term precipitation analyses. Vogel and Fennessey (1993) have shown similar results and provide an analytically simple algorithm for calculating L-moments from a data set, and Hosking (1989) demonstrates how they can be used to calculate the parameters and quantiles of the underlying distributions.

Forecasting and Prediction

One important aim of modeling is to understand how current and past observations will lead toward the future state of a system. It is important to note, however, that the means by which a forecaster achieves this aim using a statistical model will often differ significantly with those of research which seek an explanation of the underlying physical processes. In many cases, it is not required that a forecast variable have a causal linkage to the variable being forecast, only that some degree of statistical correlation between the predictor and the predictand be established. Such models are referred to prediction models.

A broad view of the literature casts most prediction techniques into two categories. The first rests on an understanding of the statistical and harmonic character of the time series of the predictand itself, without using (in a direct sense) any other variables to aid in the prediction process. Examples of such techniques include harmonic and spectral analysis (including Fourier analysis and singular spectral analysis) (e.g.,

Mitra et al., 1991), and modeling the time series as an autoregressive process (Wilks, 1995, page 306). These methods presume that all of the information needed to make viable predictions is contained within the predictand time series itself, since ultimately the predictand represents the aggregate of all processes which contribute to its formation, and therefore adequate predictions can be created without resorting to other information sequences.

The second category attempts to mitigate these problems by employing predictor variables, which presumably have some relationship with the predictand (although this relationship need not be causal). A simple linear regression is an example, with some variable (or set of variables) expressing a linear relationship with the variable being predicted. This system can be easily extended to include non-linear relationships between the predictor variables and the predictand.

A serious limitation of basic regression techniques is that the same set of predictors is used for the entire time series of the predictand. Thus, if one set of predictors will only fit a subset of the predictand time series, the remainder of the series will continue to be affected by its presence in the regression model. An alternative to standard regression techniques is using polynomial-based splines (De Boor, 1978; Friedman and Silverman, 1989).

A *spline* is simply a function that is used to fit only a portion of the data series being modeled. Thus, each set of splines being used to approximate the behavior of some sequence of data act as “basis functions” (functions which are used to form the basic equations for the fitted model) only for that portion of the data being modeled, and each set of basis functions are piecewise continuous with all other basis function sets which

span the predictand series (Friedman and Silverman, 1989). These individual basis functions can be of any type which fit the needs of the model, but the most common type are linear and cubic splines; cubic splines are especially popular due to their desirable quality that they are differentiable. The spline sets are joined together in a piecewise-continuous fashion such that the entire predictand series is modeled; the joining locations are known as *knots* (Friedman and Silverman, 1989; Friedman, 1991).

Splines are problematic due to their piecewise and possible discontinuous nature. However, Friedman (1991) has published an algorithm which outlines the procedure by which a spline set can be tested for optimal fit and parsimony, which he has referred to as Multivariate Adaptive Regression Splines, or MARS. This technique works through the successive testing and eliminating of the spline basis functions, and then refitting the knot locations by regressing the results at each stage of the process and comparing with the original predictand series. In this way, nearly every combination of predictors is tested, and only those that reduce the overall piecewise error are kept. The error is measured using the generalized cross validation (GCV) criterion introduced by Craven and Wahba (1979), which is output by the MARS procedure. This algorithm has been employed by Lall et al. (1996) and Sangoyomi et al. (1996) to examine the non-linear characteristics of the depth and volume variations of the Great Salt Lake, and by Lall and Mann (1995) to measure low-frequency climate variability in the Great Basin region of the U.S.

El Niño - Southern Oscillation

Few areas of climatic research have garnered as much attention in recent years as the El Niño-Southern Oscillation (ENSO) phenomenon. For many researchers, ENSO represents the epitome of ocean-atmosphere couplings, and its effects reach far beyond

the tropics. With greater varieties of data sets and more advanced observational and computational tools at their disposal, climatologists are now better able to comprehend the underlying mechanisms surrounding the ENSO, and have a clearer understanding of the tropical/extratropical ocean-atmosphere connections that influence the earth's climates.

The periodic appearance of anomalous precipitation in northwestern South America has been recognized for centuries, and first appears in the literature in the 1800s. It was noted that the precipitation arrival often coincided with Christmas, and the colonists christened the event "El Niño," after the Christ child (Enfield, 1992); the term "El Niño" would later be used to refer to the warm ocean currents associated with the anomalous precipitation. While the increased rains were often welcomed in the otherwise arid region, the rainfall sometimes reached such extreme levels that flooding occurred, causing farmland and structures to be destroyed. In addition, there were massive periodic fish and bird deaths that were often associated with these events (Sharp, 1992).

In the early twentieth century, another atmospheric phenomenon was discovered, which was to be closely related to the El Niño. Sir Gilbert Walker, a British statistician stationed in India, noted an inverse correlation between the pressure fields in Indonesia and the central Pacific Ocean (Walker and Bliss, 1932). This pressure relationship became known as the "Southern Oscillation," noting its tendency to shift phases in a quasi-periodic manner. The high and low pressure regions would modify over a period of two to five years on average (Rasmussen and Carpenter, 1983), roughly the same frequency of change associated with the El Niño, as pointed out by Berlage (1957). Nevertheless, many atmospheric scientists were reluctant to couple the two phenomena without firm evidence that they were related, especially since not all oscillation swings were tied to El Niño events (Deser and Wallace, 1987; Enfield, 1989).

While the physical processes may not have been understood, many of the manifestations of these processes certainly were, and this led to the development of measures of the severity of the El Niño and the Southern Oscillation. One measure of the pressure gradient between the central Pacific Ocean and the Indonesian archipelago was the Southern Oscillation Index (SOI), which represented the normalized sea level pressure differences between Tahiti and Darwin, Australia (Caviedes, 1975, 1984; Diaz and Pulwarty, 1992). Relatively strong negative values of this index were associated with low east-west pressure gradients in the equatorial Pacific, and weak easterly Trade winds, while strong positive values were associated with stronger Trade winds (Diaz and Kiladis, 1992). The term “El Niño” has come to refer to the former condition, while the latter is variously referred to as “anti-El Niño,” “La Niña,” and “El Viejo.” (For purposes of brevity, the term “La Niña” will be used in this document to identify the opposing conditions to El Niño.)

The connection to wind fields was important, because it was well understood that upwelling was caused by wind shear across the ocean surface. Upwelling is an important aspect of ocean-atmosphere relations, both from a biological and physical standpoint. Upwelling helps to recycle deep nutrient-rich waters in several ecologically sensitive regions, notably off the coast of western South and North America (Ochoa and Gomez, 1987). Upwelling can also drive climatic conditions within a region, particularly when the upwelled waters are from below the primary thermocline; cold upwelling can create domes of cool air, which have a stabilizing effect on the local meteorology, and thus lead to regions of relative aridity (Steiner and Khalsa, 1987). This is especially apparent in the coastal regions of Peru, Ecuador, western Central America, and California (Umatani and Yamagata, 1991).

In the 1960s, Jacob Bjerknes began formulating ideas that attempted to link the El Niño phenomenon with the Southern Oscillation using internal dynamic mechanisms,

rather than phenomenologically as with previous research. His arguments linked the local El Niño phenomenon experienced in South America with the global relaxation of the prevailing Trade winds in the equatorial Pacific (Bjerknes, 1966, 1969). Bjerknes believed that the El Niño effects were the direct result of a similar local relaxation of the winds offshore of South America. While this was later shown to be incorrect, his ideas laid the foundation for most future research into ENSO (Enfield, 1989).

Bjerknes was no doubt hampered by the lack of high-resolution observational data, which became available after 1975, due in part to increased research interest in the ENSO after the intense 1972 El Niño episode. It was found that, contrary to common assumptions, the winds offshore Peru did not slacken along with the remainder of the equatorial Trades during low phase periods of the SOI (Enfield, 1989). This forced a reconsideration of Bjerknes' theories concerning the presence of warm waters offshore of South America during El Niño periods. What has evolved instead is a much more complex model of ocean-atmosphere interaction.

The application of wave dynamics to the coupled ocean-atmosphere system has given rise to the most successful theories regarding the ENSO phenomena, and demonstrates how the atmospheric conditions of the western equatorial Pacific can directly impact on the oceanic conditions off the coast of South America. The Trade winds in the vicinity of the equator generally exercise enough shear on the surface waters to force them westward, thereby causing a rise in sea level in the west, and a commensurate depression of the primary thermocline (Philander, 1989). During periods when the Trades slacken, the warm surface waters in the west would begin to move eastward. Due to the Coriolis effect, water north and south of the equator is forced towards the equator, thus entraining the eastward propagating wave of water. This is known as a Kelvin wave (after Lord Kelvin, who is credited for having first described them mathematically) (Enfield, 1989).

When the wave encounters the east continental boundary, three things can happen to the energy being carried with it. First, a certain amount of the energy is reflected directly back towards the west; in this situation, the Coriolis effect works against the wave train, driving northern water more towards the north and southern waters more towards the south, thus quickly dispersing the wave (Cane and Gent, 1977; Horel and Wallace, 1981; Knox and Halpern, 1982). Second, much of the energy is transferred into westward propagating counter-rotating eddies called Rossby waves (Horel and Wallace, 1981; Keen, 1982). Third, the sea levels are pushed upward (and the thermocline depressed) along the continental boundary; it is this effect which causes the temperature of the waters along northwest South America to rise during an El Niño event, since the upwelling now comes from warmer waters having been advected from the west (Wyrtki, 1975, 1984; Harrison and Schopf, 1984).

The shifting of the equatorial waters is responsible for many of the atmospheric phenomena that are associated with ENSO. During a "normal" phase of the ENSO (i.e., an SOI near the long term average), the warm waters held in check by the Trade winds in the western Pacific destabilize the atmosphere over the region, causing widespread convection due to warm air over the warm water. In the east, cold waters near South America stabilize the atmosphere, causing subsidence and precluding widespread precipitation. The subsiding air in the east is carried westward by the trade winds, lifted, and carried back towards the east, completing the so-called "Walker circulation" of air over the equatorial Pacific (Julian and Chervin, 1978; Rasmussen and Carpenter, 1983). (Similar atmospheric circulation features have been noted over the Indian and Atlantic Oceans.) Thus, the prevailing arid conditions over Peru and Ecuador, as well as the wet conditions over Indonesia, can in large measure be attributed to the coupling of the prevailing winds and the advection of warm equatorial waters of the Pacific Ocean (Gill and Rasmussen, 1983; Goldberg et al., 1987).

If the pressure gradient is reduced, causing a relaxation in the Trades and the triggering the onset of an El Niño, the migration of warm waters causes the areas of maximum convection to be shifted towards the east (Philander, 1981). This disrupts the Walker circulation, causing subsiding conditions to move over Indonesia, often resulting in drought conditions because of colder than normal ocean waters (Julian and Chervin, 1978). In the east, warm waters destabilize the air over Peru and Ecuador, sometimes causing intense rains and flooding.

The ENSO effects extend well beyond the tropics. Through the process of *teleconnection*, semi-permanent pressure centers in the mid- and high-latitudes can be shifted, causing dramatic changes in the mid-latitude jet stream patterns, and resulting in sudden and significant changes to the temperature and precipitation patterns of regions well outside the tropics (Van Loon and Madden, 1981). The primary reason is that the zonal shortening of the Walker circulation associated with ENSO forces a commensurate meridional increase in the Hadley circulation, which is responsible for variations in the Ferrel and Polar circulation cells, and therefore the polar and subtropical jet streams as well (Horel and Wallace, 1981). These changes in the jets are responsible for variations in mid-latitude weather system movement. For example, an eastward shift in, or a deepening of, the Aleutian Low by just a few hundred miles can result in winter storms impacting the California coast far to the south of their expected strike zone (Emory and Hamilton, 1985); this occurred during the 1982-83 El Niño episode (Baumgartner and Christensen, 1985; Simpson, 1984). A shift in the opposite direction can cause extreme drought in the Southwest United States. Changes in the mean storm track into the west coast of the U.S. is often reflected in the so-called Pacific-North American (PNA) index, which compares 500 millibar geopotential heights for locations in the north-central Pacific Ocean, the southeastern U.S., and the northern Great Plains of North America; a positive correlation between the PNA and the SOI has been noted during the boreal

winter months, which is indicative of anomalous high-amplitude shifts in the polar jet stream, which tends to direct more storms into southern California and the Southeast U.S. during ENSO events (Epstein, 1992; Halpert and Ropelewski, 1992). A similar linkage between the ENSO and precipitation in the Southwest United States was noted by Woolhiser et al. (1993). Katz and Parlange (1993) further demonstrate the ability to condition stochastic rainfall occurrence parameters in the United States on an atmospheric index. This invites investigation into whether the daily rainfall properties of tropical Central America can be similarly conditioned on an index such as the Southern Oscillation Index (SOI), which gives an indication of the state of the ENSO.

Several studies have examined the effects of ENSO on the Central American isthmus. An examination on the relationship between precipitation and the state of ENSO by Ropelewski and Halpert (1987) and Halpert and Ropelewski (1992) demonstrated the prevalence of drought-like conditions over much of the region during El Niño events. Estoque et al. (1985) shows that El Niño periods are associated with decreased rainfall in most of Panama, while northwest Panama receives less rainfall during the same period; this supports the overall conclusions of Ropelewski and Halpert, while hinting at further complexity in the region. Further studies in Costa Rica by Fernández and Ramírez (1991) reiterated these conclusions, and noting a connection between the Trade winds and the west-to-east contrast in the rainfall characteristics. Further research by Waylen et al. (1996) showed that the pattern of ENSO effects on rainfall in Costa Rica was more complex, with increased precipitation during El Niño events occurring in the eastern part of the country, and thus opening the door to further speculation on the nature of ENSO-based rainfall modification in the region.

CHAPTER 3

STUDY AREA AND DATA DESCRIPTION

The primary physiographic features and climatic processes which influence the formation of precipitation and facilitate the interpretation of daily precipitation models in Costa Rica are discussed. The study area is delineated, meteorological stations are identified, and a subset of three stations is selected for examination. The amount of daily rainfall data available from each station is determined. In addition, the climate indices used in the correlation analyses are identified.

Physiography

The topography of Costa Rica is a complex mixture of humid coastal lowlands to the east and semiarid plateaus to the west, separated by *cordilleras* which roughly bisect the country from northwest to southeast, with an embedded Central Valley (Figure 3-1). Four primary *cordillera* are evident. The northern part of the country is dominated by the Cordillera Guanacaste, Tilaran, and Central, which are volcanic in nature; and in the south by the Cordillera de Talamanca, which represents an exposed batholith. Elevations range from sea level to over 3000 meters in the cordilleras, with the greatest elevation found in the south. The cordillera are punctuated with narrow passes, which allow dominant winds from one side of the country to penetrate to the lee side; these passes are more numerous in the northern cordillera, which have lower elevations and less east-west extent than the southern cordillera.

General Climatic Characteristics

Costa Rica lies at the confluence of several weather-generating processes, which give rise to the seasonal precipitation patterns seen in Figures 3-2 and 3-3. The region is dominated by the northeast Trade winds, which keep the eastern lowland well-watered, but result in a rainshadow to the west. The Trade winds reach their greatest strength during the boreal summer months, when the pressure gradient between the Atlantic Subtropical High and the ITCZ is maximized. This low-level flow results in the advection of an large amount of precipitable water from the Caribbean, which is condensed when the moist air is forced over the cordillera. The ITCZ experiences a north-south migration in the eastern Equatorial Pacific, from approximately 3 degrees north in the boreal winter to 10 degrees north in the summer. The ITCZ promotes considerable instability in the region, bringing heavy rains to the western areas of Central America. In addition, cross-equatorial westerlies arise south of the ITCZ, bringing frequent precipitation to the extreme south of the country as the ITCZ reaches its maximum northerly extent (Hastenrath, 1976). It is also during the northward transit of the ITCZ that a strengthening of the Trade wind flow occurs, resulting in the “Veranillos de San Juan” (or simply *veranillos*, literally, “little summer”), during which the rainfall abates somewhat; this generally occurs during July, and ends when the ITCZ comes to its most northerly position in late summer.

During the boreal summer, there has also been reported a connection between the occurrence of tropical storm activity in the Caribbean basin and rainfall on the Pacific coast of Costa Rica; it has been posited by Vargas and Trejos (1994) that this is due to a

reversal in the pressure gradient over this region during periods of tropical storm activity, altering the overall wind fields over the Central American isthmus.

A third seasonal influence is the incursion of polar air from the north during the winter months, following cold front penetration into the region; these are referred to as the *nortes*. This results periods of prolonged precipitation in the northeastern part of the country, a consequence of stress differential induced convergence by the northerly winds (Bryson and Khun, 1961; Schultz et al., 1997, 1998). While less effective at creating precipitation over long periods than the summer climate components, the *nortes* are nevertheless important rainfall components, since the rains they produce come at a time when the country is experiencing considerably drier conditions than in the summer. The *norte*-driven rainfalls are often locally quite intense, resulting in the unusual situation where the annual maximum rainfalls for the area will occur during the winter months.

Previously Noted Costa Rica-ENSO Connections

The convergence of these processes makes the climate of Costa Rica extremely responsive to changes in circulation in many parts of the global climate system; Waylen et al. (1996) has reported significant correlations between monthly rainfall and SOI measures with lags of up to +/- two years, demonstrating the long-term sensitivity of the region's climate to changes in global circulation.

Extremes in the state of the ENSO, as measured by the SOI, have been shown to affect the climatic conditions over the Central American isthmus. El Niño conditions tend to enhance the subtropical jet stream to the north during the boreal winter, while increasing the Trade winds in the south over Costa Rica and Panama; this is due to the increased pressure gradient that exists between the equatorward north Atlantic high

pressure system and the lower pressures associated with the warm waters off of Peru and Ecuador (Estoque et al., 1985; Fernández and Ramírez, 1991; Waylen et al., 1994, 1996). During low phase summers, northward movement of the Pacific ITCZ is restrained, causing significant precipitation deficits in northwest Costa Rica (Waylen et al., 1996). In the Atlantic during the boreal summer, increased Trade winds arise due to anomalous movement of the ITCZ, resulting in a decrease in the incidence of tropical storm activity (Vargas and Trejos, 1994).

During La Niña conditions, the northward migration of the ITCZ during the summer allows increased precipitation to the west coast of Central America, and reduced Trade wind activity over the Atlantic (Waylen et al., 1996); these periods are often associated with increased Atlantic tropical storm activity (Gray, 1984a,b), and greater Westerly winds along the Pacific coast. During the winter of La Niña, an increased meridional pressure gradient between North and Central America allows greater cold front intrusion into Nicaragua and Costa Rica, resulting in more intense *nortes*, and increased rainfall in the northeast regions (Schultz et al., 1997, 1998).

In terms of regional impacts, the progression of a typical ("canonical") El Niño event is fairly well documented. During the year of the onset of ENSO ("Year 0"), the waters offshore of northwestern South America begin to warm beyond the seasonal averages, the result of the advected waters associated with the Kelvin wave moving eastward in the equatorial Pacific. The warm waters result in lower pressures in the eastern Pacific, which intensifies the pressure gradient between the Azores high pressure in the Atlantic and the eastern Pacific, bringing an increase in the Trade winds across the Central American isthmus. The Trade winds increase the rainfall on the eastern slope of

Costa Rica, and increase the upwelling associated with the Costa Rica Dome, an area of colder waters offshore of western Costa Rica. This causes a more stabilized atmosphere to develop over the western half of the region; the ultimate result is reduced summer precipitation in the west, and higher than normal summer rainfall in the east. Increased wind stress over the Caribbean (resulting from increased Trade winds) causes cooler water to surface, raising the atmospheric pressure in the region and reducing the overall pressure gradient between the North American continental high and the Caribbean low. The lower pressure gradient results in a reduced capacity for cold air masses to move southward, causing fewer *norte* events to occur during ENSO year 0. Figures 3-4 and 3-5 show how Year 0 seasonal rainfall differs from the long-term average.

In the year following an ENSO event ("Year 1"), the climatological impacts in the region are reversed from the Year 0 conditions. This reflects the reestablishment and strengthening of the ITCZ, the increasing Trade winds over the central Pacific equatorial waters, and the resulting westward advection of warm waters (Figures 3-6 and 3-7). The resumption of cold conditions offshore of South America reduced the pressure gradient between the North Atlantic High and the western Pacific, resulting in an overall decrease in the Trade wind activity across Central America, and a commensurate reduction in summer rainfall in the eastern part of the country, as well as an increase in convective activity in the west. Additionally, *norte* activity increases during the winter months of Year 1, probably resulting from increase latent heat being drawn into the Caribbean basin as a consequence of a more northerly ITCZ during this time; this helps create an increased pressure gradient between the wintertime North American high pressure and the low pressure over the Caribbean, which allows greater cold front activity farther

south. The increase in *norte* activity bring increased Year 1 rains to the northeastern part of the country. Schultz et al. (1997, 1998) hypothesize a differentiation between the strength of the *nortes* between Year 0 and Year 1, referring to them as “cool” and “cold” *nortes*, respectively, based upon the strength of the cold air masses as they are modified by the El Niño.

Daily Precipitation Data and Station Location

The greatest difficulty faced in performing a detailed climatic analysis is the acquisition of reliable, long-term data. This is especially true when the analysis entails the use of daily precipitation data. In many cases, such records were not kept in their raw state, but rather were aggregated into monthly and annual records, rendering them useless for this particular study. In Costa Rica, however, many of the original daily rainfall measurements exist in digital form, and they are considered a commercially valuable commodity by the agencies that control them.

For reasons of expense, it is important to narrow the focus for data acquisition to that which is absolutely necessary for the completion of the study. In this case, the focus is on describing a transect which encompasses the most populous (and thus most densely monitored) regions of the Central Valley, stretching eastward from the Caribbean coastline (enclosing an area approximated by the drainage basin of the Reventazón River), and westward to the Pacific Ocean (to include the basin of the Tarcolés River). There are 41 stations available within this transect; however, only eight stations possess at least 30 years of records. It is from these eight stations, which places significant limitations on the number of stations that can be used for long-term analysis. Table 3-1 lists the these stations, and Table 3-2 shows their geographic disposition and data

availability. Figure 3-8 shows the locations of the stations, and the boundary of the study area.

Three stations were chosen from the eight long-term stations (those with records in excess of 30 years) to be used for performing detailed statistical analyses. The stations were chosen to be representative of the Caribbean, Central Valley, and Pacific coast regions of the study area, because of their geographic locations and relatively long data records. Sanatorio Duran (station ID number 073011) was selected to represent the Caribbean region, Villa Mills (073033) was chosen for the Central Valley/transition region, and Nagatac (080005) was chosen to represent the Pacific coast. These stations are highlighted on Figure 3-8. The monthly rainfall characteristics of these three stations in relation to ENSO is shown in Figures 3-9 to 3-11.

Atmospheric Index Data

Atmospheric index data, provided by NOAA via the World Wide Web (NOAA, 1996) was used to perform lag-correlations on the computed daily rainfall statistics. Four indices were selected to represent the likely areas of broad-scale climate impact upon the Costa Rican rainfall generation process. The indices selected were the Southern Oscillation Index (SOI), the North Atlantic Oscillation index (NAO), the central equatorial Pacific sea-surface temperature index (NINO34), and the Quasi-Biennial Oscillation index (QBO50).

The SOI is calculated by subtracting the normalized monthly sea level pressure (SLP) at Darwin, Australia from the normalized SLP at Tahiti; the SOI provides a reasonable representation of the overall pressure gradient that exists between the central and western Pacific regions, and is the traditional means by which the strength of an El

Niño event is measured. The NAO is calculated in the same fashion as SOI, substituting the SLP values for the Azores and Iceland; the NAO measures the strength of the SLP dipole over the North Atlantic, and as such might give an indication into the role of the North Atlantic high pressure zone in the development of Trade winds across Costa Rica. In the case of both the SOI and the NAO, each monthly measurement is normalized by subtracting the average SLP for a month from the long-term average SLPs at that site, and dividing by the standard deviation of the long-term SLPs. The data for both indices have been recorded continuously since the 1800s, and are readily available to any researcher.

The NINO34 index is important due to the strong association between abnormally warm ocean temperatures of the equatorial Pacific and the onset of an ENSO event. The NINO34 measurements represent a monthly averaging of the equatorial sea surface temperatures surrounding the 180 degree meridian, and are calculated by using a combination of historical shiptrack data augmented in recent years by remotely-sensed satellite data.

The QBO50 index is an average of the zonal equatorial stratospheric winds at the 50 millibar level. These winds have been shown to possess quasi-periodic cycles of approximately 27-30 months, during which time they generally change direction twice; from this cycle is derived the designation "Quasi-Biennial Oscillation" (QBO). While the physical mechanisms that govern the interaction of tropospheric and stratospheric climate systems are poorly understood, a number of researchers have noted strong statistical associations between climate anomalies and the QBO (Kodera, 1991; Mason and Tyson, 1992). Gray (1984a, 1984b) has demonstrated a correlation between both the QBO and

SOI and the formation of tropical cyclones in the Atlantic basin. Vargas and Trejos (1995) noted a possible connection between Atlantic tropical cyclone activity and Central American rainfall, and for this reason the QBO50 index was included in the analysis.

For many of the analyses it was necessary to segregate the data record into “El Niño,” “La Niña,” and “other” year-types. This segregation was based upon the classification scheme used by Diaz and Kiladis (1992) and supplemented with NOAA data to include the most recent events. The resulting year listing is shown in Table 3-3.

Table 3-1. Station locations*

Station ID Number	Station ID Code	Station Name	Latitude	Longitude	Elevation**
073011	SANDURAN	Sanatorio Durán*	9.93	-83.88	2337
073016	JUANVINA	Juan Vinas	9.83	-83.78	990
073026	TAPANTI	Tapanti	9.77	-83.83	1203
073027	ELCANON	El Cañón	9.68	-83.90	2460
073028	ELHUMO	El Humo	9.80	-83.72	680
073029	ELLANO	El Llano	9.77	-83.87	1572
073030	CORDONCI	Cordoncillal	9.75	-83.78	1240
073033	VILLAMIL	Villa Mills*	9.57	-83.72	3000
073034	BELÉN	Belén	9.73	-83.88	2010
073036	TSEIS	T-Seis	9.72	-83.77	2000
073037	ELDESTIE	El Destierro	9.70	-83.75	2020
073038	OJOREVEN	Ojo de Agua	9.62	-83.82	2960
073039	TRESDEJU	Tres de Junio	9.67	-83.85	2200
073040	BERMA	Berma	9.67	-83.82	2480
073044	LASUIZA	La Suiza	9.85	-83.62	620
073045	TAUS	Taus	9.77	-83.72	900
073046	CACHIPLA	Cachí (plantel)	9.82	-83.80	1018
073047	TUCURRIQ	Tucurrique	9.85	-83.75	770
073055	LAAMISTA	La Amistad	9.98	-83.90	560
073074	ANTONTUR	San Antonio de Turrialba	9.97	-83.72	1190
073080	LAMUERTE	Cerro La Muerte (Repet.)	9.57	-83.77	3365
073096	ELSAUCE	El Sauce	10.00	-83.65	870
080005	NAGATAC	Nagatac*	10.07	-84.55	450
084005	AVANCEDE	Avance de Tres Ríos	9.98	-83.97	1870
084006	HACONCEP	Hca Concepcion Tres Ríos	9.92	-84.00	1320
084018	REDONDO	Rancho Redondo	9.95	-83.95	1780
084019	LAGUNA	Laguna	9.97	-83.87	3140
084024	OJOTARCO	Heda Ojo de Agua	9.93	-84.22	850
084027	SANRAMON	Subestación San Ramón	10.08	-84.47	1061
084034	EMBALAGA	Embalse La Garita	9.95	-84.35	460
084037	ANTONESC	San Antonio de Escazu	9.90	-84.13	1380
084041	IPISDEGU	Ipis de Guadalupe(Planta)	9.97	-84.03	1280
084043	SDROBLE	Santa Domingo del Roble	10.07	-84.17	1320
084047	PRESALAG	Presa La Garita	9.93	-84.30	471
084048	ANTECAMA	Antecámara La Garita	9.93	-84.35	466
084063	VOLCANPO	Volcán Poas	10.18	-84.23	2564
084074	PAVAS	Pavas	9.97	-84.13	997
084103	ELDESCAN	El Descanso	10.02	-84.20	800
084104	LAMARAVI	La Maravilla	9.90	-83.77	-9.9
084105	SUBHERED	Subestacion Heredia	10.00	-84.13	-9.9
084110	LAGUACIM	La Guacima	9.97	-84.25	-9.9

* Bolded stations were used for in-depth focus in this study.

** A value of -9.9 indicates that the elevation has not been accurately surveyed.

Table 3-2. Station data lengths.

Station ID Number	Station ID Code	Station Name	Start Year	End Year	Percent Missing Data	Available Years of Record
073011	SANDURAN	Sanatorio Durán	1943	1991	2.01	48.0
073016	JUANVINA	Juan Vinas	1975	1991	0.52	16.9
073026	TAPANTI	Tapanti	1939	1991	7.38	49.1
073027	ELCANON	El Cañón	1954	1991	7.09	35.3
073028	ELHUMO	El Humo	1954	1991	4.6	36.3
073029	ELLANO	El Llano	1957	1991	4.2	33.5
073030	CORDONCI	Cordoncillal	1960	1991	9.15	29.1
073033	VILLAMIL	Villa Mills	1942	1991	6.66	46.7
073034	BELEN	Belén	1959	1991	22.03	25.7
073036	TSEIS	T-Seis	1962	1991	13.99	25.8
073037	ELDESTIE	El Destierro	1962	1991	23.11	23.1
073038	OJOREVEN	Ojo de Agua	1959	1991	10.34	29.6
073039	TRESDEJU	Tres de Junio	1962	1991	2.19	29.3
073040	BERMA	Berma	1962	1991	19.76	24.1
073044	LASUIZA	La Suiza	1962	1991	2.62	29.2
073045	TAUS	Taus	1962	1991	26.21	22.1
073046	CACHIPLA	Cachí (plantel)	1952	1991	0.16	39.9
073047	TUCURRIQ	Tucurrique	1964	1991	0.21	27.9
073055	LAAMISTA	La Amistad	1966	1991	13.03	22.6
073074	ANTONTUR	San Antonio de Turrialba	1965	1991	2.93	26.2
073080	LAMUERTE	Cerro La Muerte (Repet.)	1970	1991	3.19	21.3
073096	ELSAUCE	El Sauce	1975	1991	3.51	16.4
080005	NAGATAC	Nagatac	1959	1997	11.7	34.4
084005	AVANCEDE	Avance de Tres Ríos	1980	1984	3.29	4.8
084006	HACONCEP	Hca Concepcion Tres Ríos	1975	1984	5.81	9.4
084018	REDONDO	Rancho Redondo	1975	1984	0	10.0
084019	LAGUNA	Laguna	1975	1984	0	10.0
084024	OJOTARCO	Hcda Ojo de Agua	1980	1984	6.63	4.7
084027	SANRAMON	Subestación San Ramón	1978	1990	16.29	10.9
084034	EMBALAGA	Embalse La Garita	1963	1990	7.24	26.0
084037	ANTONESC	San Antonio de Escazu	1980	1983	0	4.0
084041	IPISDEGU	Ipis de Guadalupe(Planta)	1969	1984	1.68	15.7
084043	SDROBLE	Santa Domingo del Roble	1959	1990	4.77	30.5
084047	PRESALAG	Presa La Garita	1971	1981	0.05	11.0
084048	ANTECAMA	Antecámara La Garita	1971	1980	4.25	9.6
084063	VOLCANPO	Volcán Poas	1971	1990	8.25	18.4
084074	PAVAS	Pavas	1975	1979	10.3	4.5
084103	ELDESCAN	El Descanso	1976	1990	3.8	14.4
084104	LAMARAVI	La Maravilla	1976	1990	19.79	12.0
084105	SUBHERED	Subestación Heredia	1976	1982	20.12	5.6
084110	LAGUACIM	La Guacima	1980	1990	5.95	10.3

Table 3-3. ENSO classification of years since 1900

El Niño Years	La Niña Years	"Other" Years	
1902	1903	1900	1952
1905	1906	1901	1955
1911	1908	1904	1956
1914	1916	1907	1958
1918	1920	1909	1959
1923	1924	1910	1960
1925	1928	1912	1961
1930	1931	1913	1962
1932	1938	1915	1966
1939	1942	1917	1967
1940	1949	1919	1968
1941	1954	1921	1971
1946	1964	1922	1974
1951	1970	1926	1976
1953	1973	1927	1978
1957	1975	1929	1979
1963	1988	1933	1980
1965		1935	1981
1969		1936	1983
1972		1937	1984
1977		1943	1985
1982		1944	1987
1986		1945	1989
1991		1947	1990
1993		1948	1992
1997		1950	1994

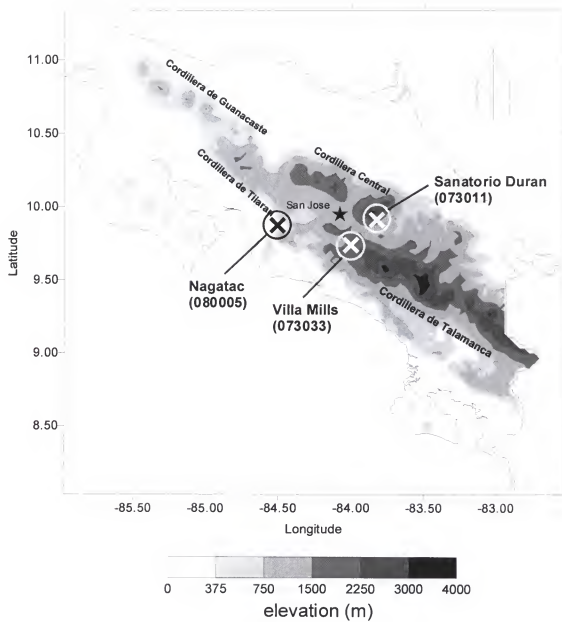


Figure 3-1. Costa Rica topography.

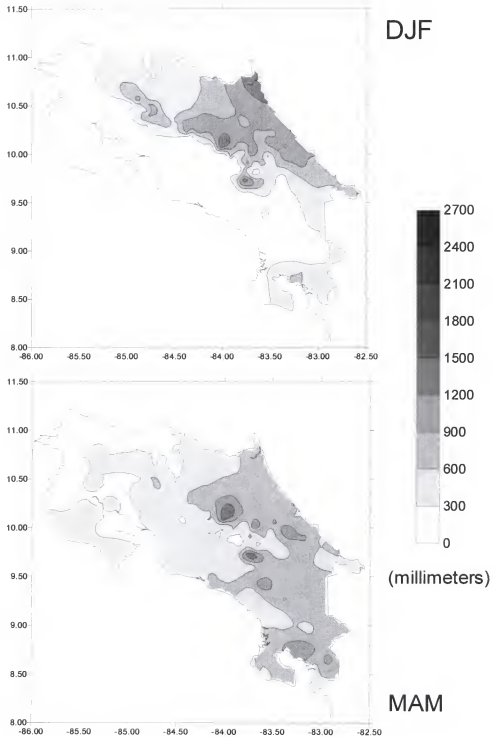


Figure 3-2. Costa Rica seasonal precipitation
Top: Boreal winter; Bottom: Boreal spring.

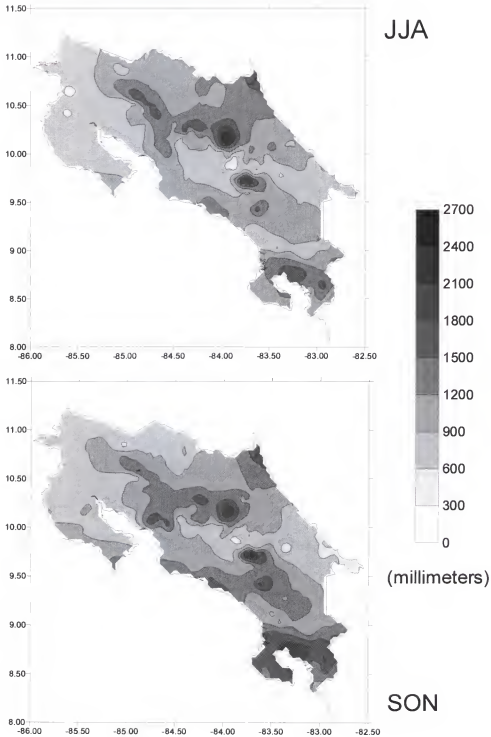


Figure 3-3. Costa Rica seasonal precipitation
Top: Boreal summer; Bottom: Boreal autumn

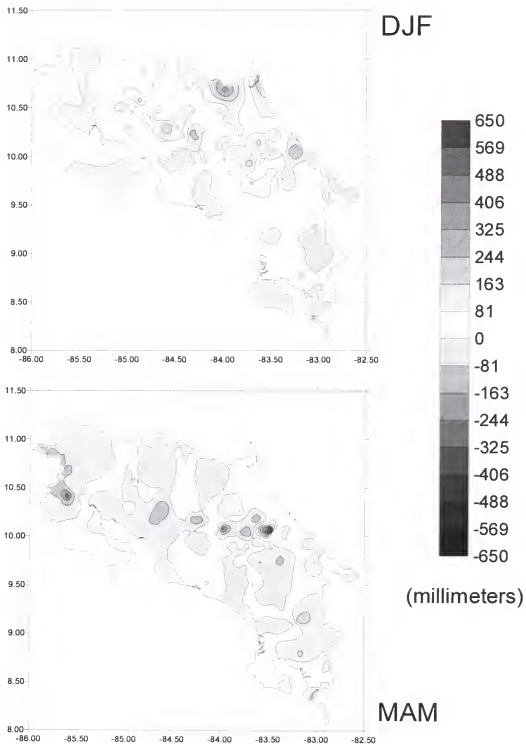


Figure 3-4. Costa Rica seasonal precipitation changes, average - El Niño years, Year 0
Top: Boreal winter; Bottom: Boreal spring

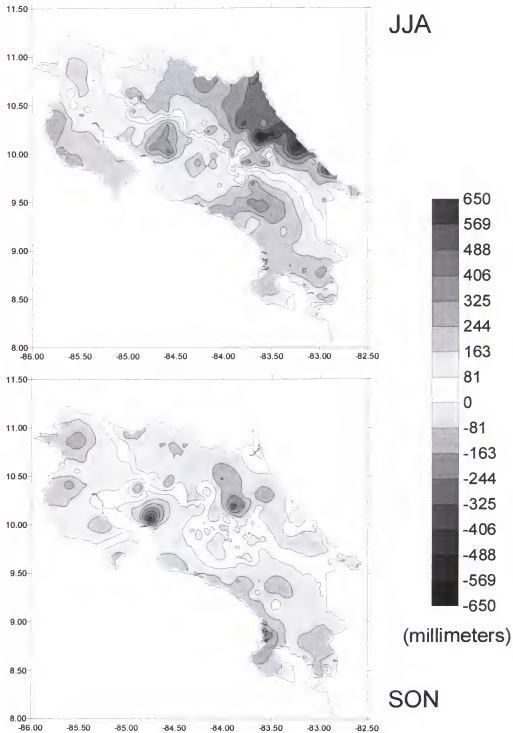


Figure 3-5. Costa Rica seasonal precipitation changes, average - El Niño years, Year 0
Top: Boreal summer; Bottom: Boreal autumn

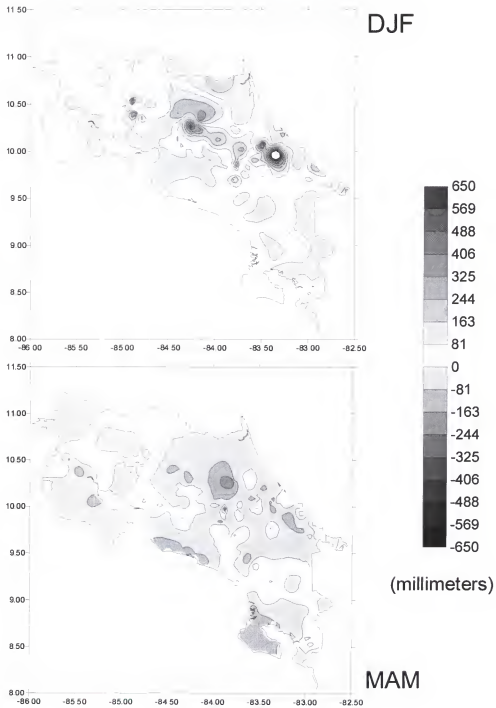


Figure 3-6. Costa Rica seasonal precipitation changes, average - El Niño years, Year 1
 Top: Boreal winter; Bottom: Boreal spring

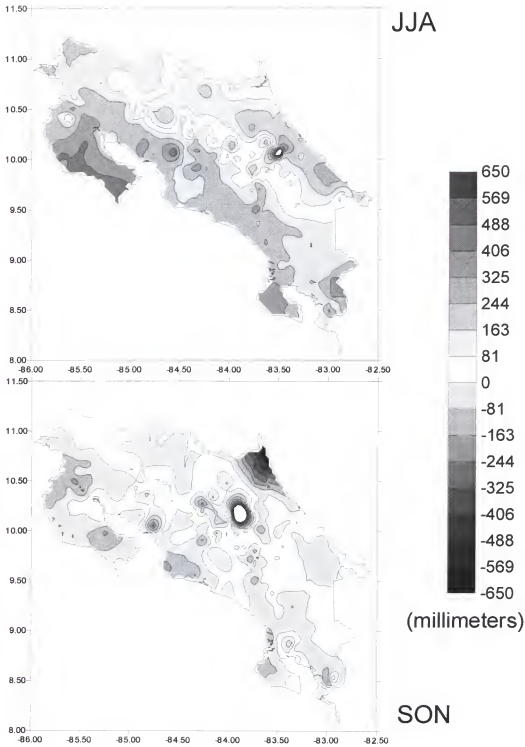


Figure 3-7. Costa Rica seasonal precipitation changes, average - El Niño years, Year 1
Top: Boreal summer; Bottom: Boreal autumn

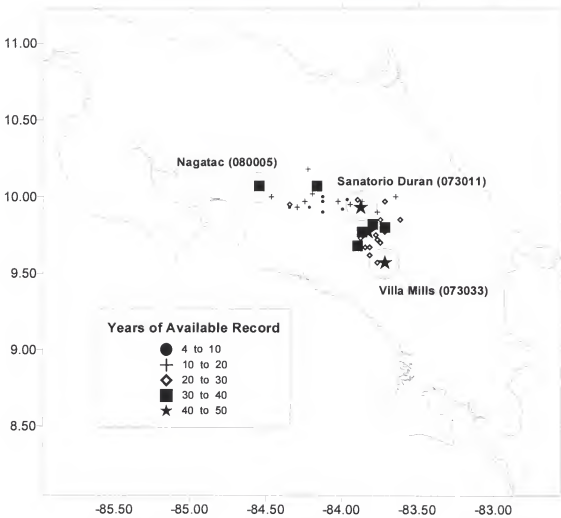


Figure 3-8. Costa Rica station location.

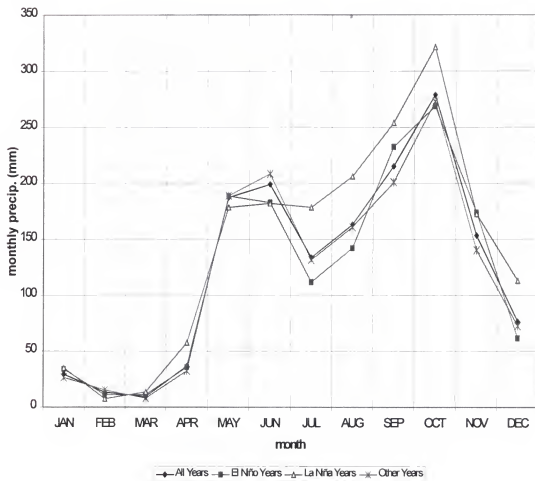


Figure 3-9. Sanatorio Duran average monthly precipitation.

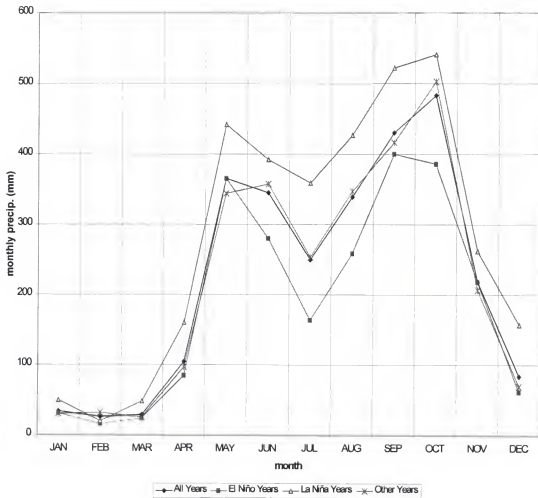


Figure 3-10. Villa Mills average monthly precipitation.

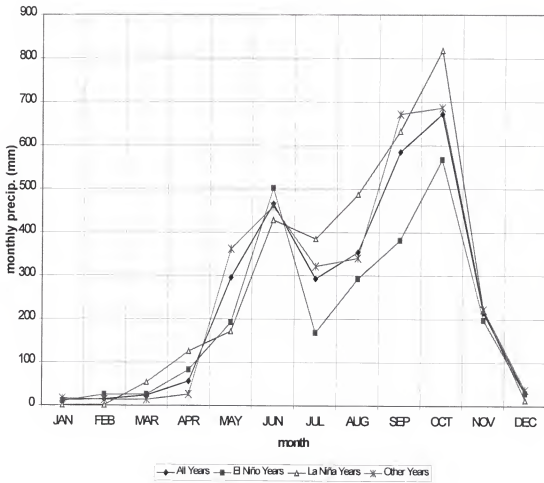


Figure 3-11. Nagatac average monthly precipitation.

CHAPTER 4 METHODOLOGY

Modeling of the occurrence and intensity of rainfall permits the creation of a comprehensive representation of daily precipitation across a variety of spatial and temporal scales. In addition, examination of the parameters describing each component facilitates the statistical correlation of model characteristics with indices describing the atmospheric processes related to the ENSO phenomenon. Within the context of the present research, the following questions will be addressed:

1. How can the characteristics of daily Costa Rican precipitation be modeled? In particular, can the occurrence and intensity probabilities be used effectively to describe the changes in rainfall that arise as a result of ENSO-related weather anomalies at individual locations? How well do each of the models compare to the actual precipitation measurements, both in terms of the goodness-of-fit and parsimony?
2. To what extent do the resulting models indicate changes in weather conditions as a result of ENSO?
3. Are the modeling results consistent with the current understanding of the climate mechanisms governing the region? How well do these results correspond to previous studies of rainfall in the region, especially those conducted at more aggregated (i.e., annual) time scales?

4. Can the modeling results be used to create reasonable estimates of future rainfall characteristics? That is, can forecasts of rainfall be created using predictions of ocean and atmospheric conditions. How well do these forecast techniques compare to each other, and to simple climatology (i.e., long-term averages)?

The general strategy for performing this research is as follows: (1) create statistics representing the occurrence and magnitude of daily rainfall events, using three long-record stations representing geographical regions within the study area, which will be aggregated into a monthly time series for each station; (2) segregate the statistics by year-type, using the augmented Diaz and Kiladis (1992) classification previously discussed, and analyze the resulting patterns; (3) perform lag-correlations on each statistic for the stations against the monthly atmospheric indices previously identified, for lags of ± 36 months, to identify possible temporal relationships between the statistics and the antecedent or following atmospheric conditions; (4) use the parameters created in step (1) to perform statistical goodness-of-fit tests to determine the best-fit distributions used to describe precipitation magnitudes for each of the stations, utilizing both Chi-squared tests and bootstrap simulation experiments; and (5) use the preceding results to construct several types of forecast models, to test the feasibility of creating “real world” forecasts, for application to agricultural and hydrological models.

Precipitation Statistics

The primary difficulty in aggregating daily rainfall data into statistics which can be meaningfully compared to existing atmospheric indices is arriving at a time scale

which will minimize the amount of inter-seasonal climate change while allowing the maximum amount of information to be represented by the statistics. Generally, a monthly time scale is used; however, this allows a considerable amount of “noise” into the transition probability measurements. For this reason, monthly triads, represented by the middle month, are computed per year of data, with each triad overlapping the previous and following triad by one month. The triad statistics are then used as a monthly time series for each of the occurrence and magnitude statistics, and can be considered “smoothed” in the sense that inter-monthly variation has been distributed across monthly boundaries, thereby reducing the “noise” and minimizing the possibility that seasonal variations will be masked in the computation of the parameters.

Precipitation Modeling -- Occurrence

Markov chain models are created by conditioning the probability of the occurrence of rainfall on one day upon whether measurable rainfall was observed on the previous days. These are characterized by “transition probabilities”:

$$\Pr\{X_{t+1} | X_t, X_{t-1}, X_{t-2}, \dots, X_1\} = \Pr\{X_{t+1} | X_t, X_{t-1}, \dots, X_{t-m}\} \quad (4-1)$$

in which the “order” of the model is given by m , the number of previous days to be examined for rainfall.

Conceivably, the number of states to be used in a Markov rainfall model is unlimited, but more generally a two-state model is used:

$$X_t = 1 \text{ wet day} \quad (4-2a)$$

$$X_t = 0 \text{ dry day} \quad (4-2b)$$

Using this notation, the transition probabilities for a two-state Markov precipitation model of order m is given by:

$$p_{ijk...m} = \Pr\{X_{t+1} = m \mid X_t = m-1, X_{t-1} = m-2, X_{t-2} = m-3, \dots, X_{t-(m-1)} = j, X_{t-m} = i\} \quad (4-3)$$

where $i, j, k, \dots$ refer to the previous days' state (i.e., wet or dry).

The maximum likelihood estimators for the transition probabilities is given by:

$$p_{hij...m} = \frac{n_{hij...m}}{n_h} \quad h, i, j \dots m = 0, 1 \quad (4-4)$$

for a multi-order, two-state Markov model, where n represents the number of transitions which have been counted for a given time period, and $i, j, k, \dots$ represent the state of previous days upon which the transition probability is conditioned.

To decide what order to use in modeling the rainfall process, it is necessary to determine how much information is being included in a model of given order, and compare this to other model orders. Of concern is the actual amount of information about the rainfall occurrence process, as well as the number of parameters to be generated to account for this information. Ideally, we desired the maximum amount of information and the fewest parameters (i.e., greatest degree of parsimony); in other words, a balance must be struck between having too many, and too few, parameters generated. "Information criteria" are employed to make this determination. The two most common are the "Akaike Information Criteria" (AIC) (Akaike, 1974) and the "Bayesian Information Criteria" (BIC) (Schwartz, 1978). Both methods rely on the generation of log-likelihood functions for the estimated transition probabilities of the fitted Markov model. Given a two-state Markov model, the log-likelihoods for orders 0, 1, 2, and 3 are given by Wilks (1995, page 301) as:

$$L_0 = \sum_{j=0}^1 n_j \ln(p_j) \quad (4-5a)$$

$$L_1 = \sum_{i=0}^1 \sum_{j=0}^1 n_{ij} \ln(p_{ij}) \quad (4-5b)$$

$$L_2 = \sum_{h=0}^1 \sum_{i=0}^1 \sum_{j=0}^1 n_{hij} \ln(p_{hij}) \quad (4-5c)$$

$$L_3 = \sum_{g=0}^1 \sum_{h=0}^1 \sum_{i=0}^1 \sum_{j=0}^1 n_{ghij} \ln(p_{ghij}) \quad (4-5d)$$

These log-likelihoods are then used to calculate the AIC and BIC scores; the order giving the lowest score is the most parsimonious (i.e., the most “efficient”) model. The criteria are calculated using the following equations for an m -order, s -state model:

$$AIC(m) = -2L_m + 2s^m(s-1) \quad (4-6a)$$

$$BIC(m) = -2L_m + s^m(\ln(n)) \quad (4-6b)$$

Given that this study will use two-state (“rain/no rain”) models, these equations simplify to:

$$AIC(m) = -2L_m + 2^{m+1} \quad (4-7a)$$

$$BIC(m) = -2L_m + 2^m(\ln(n)) \quad (4-7b)$$

The BIC is generally considered the more conservative of the two criteria, and will often give more definitive results than the AIC. In this study, both AIC and BIC are calculated and compared, in order to determine the best order to be used for the Costa Rica data.

Most studies of daily rainfall (e.g., Katz, 1977, 1981; Coe and Stern, 1982) have found that a first-order model is sufficient to describe daily rainfall. This allows the

occurrence process to be represented by four parameters, of which two must be calculated:

$$p_{ij} = \Pr\{X_j = n \mid X_i = m\} \quad m, n = \{0, 1\} \quad (4-8)$$

where

$$X_n = \begin{cases} 1 & \text{if precipitation occurred} \\ 0 & \text{otherwise} \end{cases} \quad (4-9)$$

and

$$p_{00} + p_{10} = 1 \quad (4-10a)$$

$$p_{10} + p_{11} = 1 \quad (4-10b)$$

The process of computing the occurrence statistics is illustrated in Figure 4-1.

Each three-month sequence is passed through a computer program which tallies the sequences of wet and dry days. These totals are used to create the transition probabilities, for orders 0, 1, 2, and 3. The AIC and BIC statistics are computed, and the resulting monthly time series of probabilities and information criteria are segregated according to whether the year-type was El Niño, La Niña, or neither. In addition, a tally is kept of all years, to be used as a baseline for comparison. An evaluation of the order results is made, and the order with the best (i.e. lowest) overall AIC/BIC is used for later analyses. In the case of mixed results, the higher order is chosen so as not to inadvertently hamper further analysis with loss of information.

Precipitation Modeling -- Magnitudes

Determination of the proper distributional form for a set of rainfall measurements is more problematic than modeling the occurrences. It is a computationally more intensive process, and sensitive to the available sample size due to the presence of low-probability high-intensity events (Guttman, 1994). In addition, since it is very unlikely

that a particular distribution will perfectly model the entire data record (both temporally and spatially), some means of choosing the best-fit distribution from among the candidates must be found.

Previous research (Hosking and Wallis, 1987 ; Hosking, 1989; Harrison, 1995) has indicated that the use of L-moments is more effective than standard method-of-moments and maximum-likelihood techniques for deriving the parameters associated with probability distributions, due to their demonstrated robustness and ease of computation. For this reason, L-moments are used in this research. The algorithms discussed in Hosking (1989) and Vogel and Fennessey (1993) are used to compute the L-moments, using the same temporal data streams as were employed in determining the transition probabilities. The L-moments are composed of “probability weighted moments” (PWM). The r^{th} PWM, b_r , which are computed using a descending-sorted data series $x_{(j)}$ composed of n elements, as follows:

$$b_r = \frac{1}{n} \sum_{j=1}^{n-r} \left[\frac{\binom{n-j}{r}}{\binom{n-1}{r}} \right] x_{(j)} \quad (4-11)$$

The L-moments, λ_r , are computed from the PWM as follows:

$$\lambda_{r+1} = \sum_{k=0}^r b_k (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \quad (4-12)$$

Analogous to product-moment ratios, Hoskings (1990) defines the L-moment ratios:

$$\tau_2 = \frac{\lambda_2}{\lambda_1} \equiv \text{L-coefficient of variation (L-CoV)} \quad (4-13a)$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \equiv \text{L-coefficient of skew (L-skew)} \quad (4-13b)$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} \equiv \text{L-coefficient of kurtosis (L-kurtosis)} \quad (4-13c)$$

The probability density functions $f(x)$, probability distribution functions $F(x)$, and inverse distribution functions $x(F)$ used in this study are listed in Table 4-1, along with the L-moment parameter estimations (Hosking, 1989).

Once computed, the L-moments are used for two purposes:

1. To select the proper distributional form that best represents the data. In the interest of parsimony, the candidate distribution is limited to those that have five or fewer parameters, but also have the ability to represent large values, such as those found in rainfall records. Candidate distributions include the exponential, generalized normal ("log-normal"), Pareto, generalized extreme value (GEV), logistic, Gumbel, gamma, Wakeby, lambda, and Pearson type III. These distributions have been extensively studied with regard to their L-moment properties, and have been widely applied to climate and hydrology problems.
2. To correlate with the atmospheric indices (e.g., SOI, NAO, Niño SST's) that describe aspects of the ocean-atmosphere system, with an aim towards understanding how ENSO affects the statistics defining the rainfall models.

The following steps will be used to determine the optimal distributional form for the daily rainfall intensities:

1. The candidate distributions (eleven total) are evaluated using a standard Chi-squared goodness-of-fit test, using five partitions (yielding 4 degrees of freedom). The results

will be used to reduce the number of candidates to approximately half the original number.

2. Simulations employing the remaining distributions and the estimated p_{xx} parameters are run, and the resulting monthly rainfall totals recorded. The “best fit” distributions are to be selected based upon a visual comparison to the historic monthly totals. The simulations are run 1000 times to insure an adequate statistical sample.
3. Further distinction between the remaining candidate distributions can be obtained by running a “bootstrap simulation” in which the five highest magnitude rain events from each of the three test stations are compared to the simulated maximum rainfalls. This method is described in detail in Wilks (1993). The simulation is run 500 times for each distribution, for each station, and for El Niño and non-El Niño years. The purpose of this test is to examine the ability of each remaining candidate distribution to accurately simulate the expected extreme rainfalls at the sites. The results will be plotted as “box-whisker” plots for two critical months, January and July, for each of the remaining distributions. The “best fit” distribution will be chosen on the basis of two criteria. First, it is desired that the actual rainfall value will closely correspond to the median value of the simulated results for each of the five largest magnitude events. Second, the variation displayed by the simulations (as illustrated by the size of the boxes and whiskers) should be as small as possible; this is indicative of the sampling variation of the simulated values (Wilks, 1993). The distribution with the highest correspondence between the simulated and actual values, coupled with the lowest sampling variation, is the distribution of choice.

Relationship to ENSO Processes — Segregated Analysis

The first step towards understanding the effect of ENSO on the occurrence and intensity patterns is to segregate the time series of calculated parameters into year-type. The segregation scheme is based on that described by Diaz and Kiladis (1992), and discussed in the previous chapter.

For each of the three stations, the data are segregated into El Niño, La Niña, and “other” years. Each year-type is then averaged on a monthly basis, resulting in a twelve-month sequence of months (January to December) for each year-type. In addition, the months for all years are averaged as well, to provide a baseline for comparing the individual year-types. The results are then plotted to determine whether there is a discernible difference between the types of years, as well as a geographic pattern based upon the location of the three stations.

Relationship to Ocean-Atmosphere Processes — Lag-Correlation Analysis

While the preceding analysis is useful in discriminating ENSO influences at a high level, it relies on the somewhat arbitrary division between ENSO and non-ENSO year-types, as well as forcing the imposition of the start and end of the calendar year upon a process that may not necessarily correspond to the cycle of geophysical processes. For this reason, it is desirable to perform a lag-correlation analysis on the calculated parameters against the time series of the atmospheric indices representing the processes of interest.

Each of the indices will be correlated against the monthly time series of parameters, using a standard Pearson product-moment correlation procedure. Each monthly observation can be correlated separately, and the index time series lagged by

± 36 months. A correlation coefficient and a statistical significance level for that correlation will be calculated.

Examination of the correlation results will allow detection of possible connections between ocean-atmosphere processes and the rainfall parameters, with the goal of establishing the degree of causality that may be present between the time series. This exceeds the simple “El Niño vs. non-El Niño” analysis of the preceding section, and attempts to illustrate possible causal linkages between Costa Rican rainfall and tropical ocean-atmosphere conditions. The strength of the correlations may also give an indication of the ability to forecast the rainfall parameters based on the atmospheric conditions that have already occurred. This application is dealt with in the next section.

Creation of Statistical Forecast Models of Rainfall Parameters

Two types of forecast models will be examined in this study, and will be tested by creating hindcast predictions for 1991, the last year of available daily rainfall data for Sanatorio Duran and Villa Mills. These hindcasts are compared to the historic parameters. For all stations and parameters, the two model types are evaluated on the basis of the degree of divergence from the actual data time series.

Multiple regression models. The first model type employs the results of the previous section to create multiple linear regression models for each parameter, based upon the highest levels of lag-correlation between the parameter and the four indices being used. Since the purpose is to test whether the parameters can be accurately forecast, only those correlations with synchronous or negative lags will be used in the creation of the models. The multiple regression model for n observations of a dependent variable y takes on the general form:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + u_i \quad i = 1 \dots n \quad (4-14)$$

where $\beta_0, \beta_1, \dots, \beta_p$ are constants referred to as the “model partial regression coefficients” (or simply “regression coefficients”) for p independent variables $x_1 \dots x_p$, and u_i represents the random error term (Chatterjee and Price, 1991). Individual regression models will be produced for each month, based upon a visual inspection of the lag-correlation results for each parameter.

ARIMA forecast models. The second model type creates ARIMA models based upon the time series of the individual parameters. This type of model does not employ the ocean-atmosphere indices, but rather assumes that the parameter time series contains all the needed information for accurate model creation. The basic form of the ARIMA model with K autoregressive parameters ϕ_k , and M moving average parameters θ_m , is:

$$x_{t+1} - \mu = \sum_{k=1}^K \phi_k (x_{t-k+1} - \mu) + \varepsilon_{t+1} + \sum_{m=1}^M \theta_m \varepsilon_{t-m+1} \quad (4-15)$$

where $(x_{t+1} - \mu)$ is the predicted anomaly for the variable as a departure from the mean value μ , x_{t-k+1} represents the previous K autoregressed values, and ε_{t-m+1} represents the weighted moving average over the previous m values (Wilks, 1995). Since the process of manually computing the parameters is very complex, the *Statistica* computer package will be used to create the models (Statsoft, 1995). The autoregressive and moving average parameters of the models will be determined based upon the autocorrelation analysis derived during the model-creation process.

Table 4-1. Distributions used in this study.*

Name	Density Function $f(x)$	Distribution Function $F(x)$	Inverse Distribution Function $x(F)$	L-moment Estimators**
Normal	$(2\pi)^{-1/2} \exp\{-[(x - \mu)/\sigma]^2/2\}$	$(2\pi)^{-1/2} \int_{-\infty}^x f(t) dt$	NAF	$\mu = \lambda_1$ $\sigma = \pi^{1/2} \lambda_2$
Exponential	$\alpha^{-1} \exp\{-x/\alpha\}$	$1 - \exp\{-x/\alpha\}$	$-\alpha \log(1-F)$	$\alpha = \lambda_1$
Gumbel	$\alpha^{-1} \exp\{-(x-\xi)/\alpha\} \exp\{-\exp\{-(x-\xi)/\alpha\}\}$	$\exp\{-\exp\{-(x-\xi)/\alpha\}\}$	$\xi - \alpha \log(-\log F)$	$\alpha = \lambda_2 / \log 2$ $\xi = \lambda_1 - \gamma \alpha$ $\gamma = \text{Euler's constant}$
Gamma	$[x^{\alpha-1} \exp(-x/\beta)] / \beta^{\alpha} \Gamma(\alpha)$ where $\Gamma(x) = \text{gamma function}$	$\gamma(\alpha, x/\beta) / \Gamma(\alpha)$	NAF	NAF (see Hosking (1989) for help)
Generalized Normal	$(2\pi)^{-1/2} \alpha^{-1} \exp(ky - y^2/2)$ where $y = k^{-1} \log[1 - k(x-\xi)/\alpha]$ $y = (x-\xi)/\alpha$	$(2\pi)^{-1/2} \int_{-\infty}^y f(t) dt$	NAF	NAF
Generalized Extreme Value	$\alpha^{-1} \exp\{-(1-k)y\} [\exp\{-\exp(-y)\}]$ where $y = k^{-1} \log[1 - k(x-\xi)/\alpha]$ $y = (x-\xi)/\alpha$	$\exp\{-\exp(-y)\}$	$\xi + \alpha \{1 - (-\log F)^k\} / k$ $\xi + \alpha \log(-\log F)$ $k \neq 0$ $k = 0$	$\alpha = (\lambda_2 k)^{-1} (1 - 2^{-1/k}) \Gamma(1+k)$ $\xi = \lambda_1 - \alpha \{1 - \Gamma(1+k)\} / k$ $k = 7.8590c + 2.955c^2$ where $c = [(2b_1 - b_0) / (3b_1 - b_0)] - [\log 2 / \log 3]$
Generalized Logistic	$\alpha^{-1} \exp\{-(1-k)y\} / (1 + \exp(y))^2$ where $y = k^{-1} \log[1 - k(x-\xi)/\alpha]$ $y = (x-\xi)/\alpha$	$1 / [1 + \exp(y)]$	$\xi + \alpha \{1 - \{(1-F)/F\}^k\} / k$ $\xi - \alpha \log[(1-F)/F]$ $k \neq 0$ $k = 0$	$k = -\tau_3$ $\alpha = \lambda_2 / \{\Gamma(1-k) \Gamma(1+k)\}$ $\xi = \lambda_1 - \alpha \{1 - \Gamma(1-k) \Gamma(1+k)\} / k$
Pearson Type III	$[(x-\xi)^{\alpha-1} \exp(-(x-\xi)/\beta)] / \beta^{\alpha} \Gamma(\alpha)$ where $\Gamma(x) = \text{gamma function}$	$\gamma(\alpha, x/\beta) / \Gamma(\alpha)$	NAF	NAF (see Hosking (1989) for help)
Wakeby	NAF	NAF	$\xi + \alpha \{1 - F\}^{\beta} / \beta - \gamma \{1 - (1-F)^{\beta}\} / \xi$	NAF (see Hosking (1989) for help)

* Entries for which no explicit analytical form exists are marked "NAF"; these must be solved by numerical iteration

(see Hosking (1989) for details)

** $b_n = n^{\text{th}}$ PWM

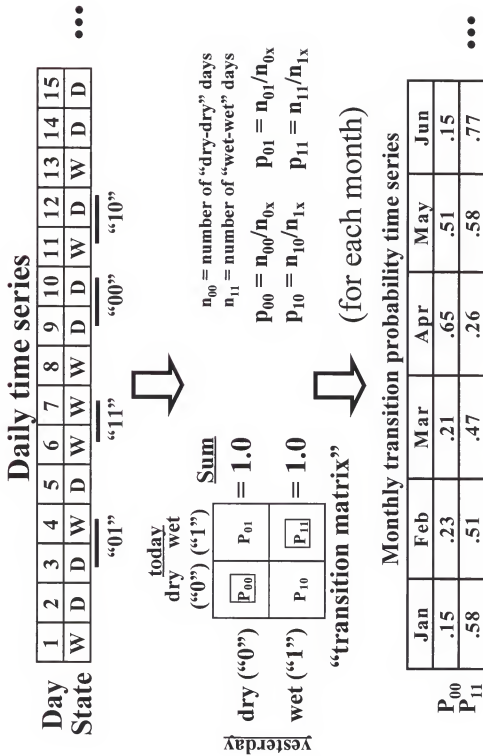


Figure 4-1. Markov model construction of daily rainfall occurrence.

CHAPTER 5

RESULTS

The identification of the proper statistical models for daily rainfall follows in a hierarchical fashion, beginning with the determination of basic model parameters and ending with the validation of the model results. This chapter will discuss the results of the model creation process, model validation, data simulation and forecasting procedures. Only the specific results are reported; more general observations and their related conclusions are discussed in the next chapter.

Model Identification

Markov Model Order

A necessary precursor to the production of Markov model parameters is the determination of the optimal order for the model. The accepted method is to examine the models using an “information criteria.” Those used are the Akaike Information Criteria (AIC) and the Bayesian Information Criteria (BIC), discussed respectively in Akaike (1974) and Schwartz (1978).

Figures 5-1 through 5-3 illustrate how frequently (as measured by the total number of station-months on the ordinate axis) a model of particular order gives the most efficient modeling results as judged by the AIC and BIC. The data were first modeled in aggregate, and then segregated into El Niño and non-El Niño years, according to the classification of Diaz and Kiladis (1992), as summarized in Table 4-4.

The graphs show a definite seasonal trend, clearly differentiating between summer and winter. During the winter, a first order model dominates using both the AIC and BIC, while summer reverts to a zero-th order model. This pattern persists through both El Niño or non-El Niño years. In all cases, the third order models are the least efficient, while the second order models in general perform less well than the first order models in terms of overall efficiency. The transition from a winter regime to a summer regime (and vice-versa) is quite rapid, requiring only a month in each case (March and November, respectively). Both the AIC and BIC indicate that the dominance of the zero-th order model rises rapidly as the summer rains resume in June, reach a maximum in October, and then falls off rapidly in early winter (November/December). The zero-th and first order models predominate over higher-order models throughout the year, with the BIC presenting a greater degree of separation between the orders, especially in the summer.

On the surface, there appears to be little to distinguish between the order of models selected in ENSO versus non-ENSO years. In fact, when viewing the winter season results on the AIC graphs, there is virtually no difference whatsoever between the segregated year types. However, the BIC tends to show a greater distinction between years in terms of order preference. During ENSO years, January becomes especially problematic, with very little to distinguish between zero-th and first orders; other winter months, however, remain strongly biased towards a first order model, using both the AIC and BIC.

The summer season is more distinctively biased towards a zero-th order model. This trend lasts for the entire length of the summer rainy season, and is especially

pronounced in the BIC. An interesting feature of both the AIC and BIC El Niño and non-El Niño graphs are the months surrounding and including July, the time of the *veranillos*. The zero-th order tally during July decreases in a subtle, but definite, fashion, which is more marked during non-El Niño years than in El Niño years.

First Order Markov Model Results – Segregated Years

The first-order Markov model parameters can be segregated based on whether the year-type is classified as El Niño, La Niña, or “other” years. Monthly estimates of p_{00} and p_{11} are shown in Figures 5-4 to 5-6, using stations Sanatorio Duran (073011, east), Villa Mills (073033, transition), and Nagatac (080005, west).

Strong seasonal trends continue to manifest themselves when examining the Markov model results for the three stations. Differences in parameter values related to ENSO are more pronounced during summer. Two notable exceptions are evident. First, Villa Mills show a distinctly higher p_{11} throughout the winter months during La Niña years, combined with a marked early winter (November/December) decrease in p_{00} . In addition, the El Niño early winter p_{00} for Villa Mills is distinctly higher than all other year types.

The separation between ENSO states is more subtle at Sanatorio Duran. During the early winter, the p_{11} for La Niña years is somewhat elevated compared to other year types, while in late winter, the trend reverses, with El Niño years becoming higher. The p_{00} parameter remains nearly the same for all year types throughout the mid- and late-winter.

Overall transition patterns. The summer patterns for all stations are more complex, due in part to the numerous precipitation-generating mechanisms that come into

play during this time. The most apparent trend is for a dramatic increase overall in the p_{11} , coupled with a synchronous decrease in p_{00} . Within this framework, however, more subtle changes can be noted. For all stations, the *veranillos* period in July/August is clearly visible as an increase in the p_{00} and a decrease in the p_{11} , giving an indication of the overall reduction of rainfall during this period. The timing and magnitude of these changes is of interest from a geographical viewpoint. The change in p_{11} at Villa Mills is slight compared to the other stations, which are located at opposite extremes of the study transect.

Transition patterns – ENSO-based variations. There is a marked summer variation in transition parameters between El Niño and La Niña years, though the separation between year types seems to be more pronounced at Nagatac and Villa Mills than at Sanatorio Duran. In addition, there are indications that the type of response has a definite geographical dependence, in that the different sides of the transect region respond in opposing fashions depending on year type.

Nagatac (west) overall shows a much higher p_{00} (and lower p_{11}) response in El Niño years during the summer compared to La Niña years, while Sanatorio Duran shows a much reduced differentiation between year types. Another feature which is clearly visible in the Nagatac p_{xx} graphs is the presence of the *veranillos* during July and August, and persisting into September; p_{00} increases during El Niño years, and the p_{11} decreases. Sanatorio Duran shows less interannual variability; the p_{xx} values are remarkably stable throughout the summer, with a somewhat greater degree of separation for p_{00} in October. Villa Mills is intermediate to the east and west stations in terms of response, but seems to show a more consistent differing response to La Niña years, with p_{00} substantially lower

during late summer, and p_{11} higher in mid-summer. One more feature of interest in the Nagatac results concerns the timing of the maximum p_{11} and minimum p_{00} values: while Figure 3-3 clearly shows that the month of maximum rainfall is October for all year types, p_{11} reaches its maximum (and p_{00} its minimum) during September for all year types except El Niño p_{00} . This would seem to indicate that the increased monthly rainfall in October must be accounted for in the magnitudes, rather than the frequencies, of daily rain events.

L-Moment Model Results – Segregated Years

The characteristics of daily rainfall magnitude (Figures 5-7 to 5-12) of the three stations under consideration correspond reasonably well to the Markov chain parameters discussed in the previous section and the total monthly precipitation characteristics illustrated in Figures 3-1 to 3-3. Increases in precipitation magnitudes seem to be associated with increased p_{11} and decreased p_{00} , as well as with an increase in total monthly precipitation.. However, several interesting deviations from this expected pattern arise, which will be discussed in this section.

Sanatorio Duran. The average daily rainfall amount (reflected in the first L-moment, L_1) for Sanatorio Duran (Figure 5-7) shows a nearly exact correspondence between total monthly precipitation and daily rainfall magnitudes. One important difference is in the early summer changes between El Niño to La Niña years. During this period (May/June), the La Niña year rains have a lower magnitude than other years, while El Niño years show higher magnitudes; during this same period, the total monthly precipitation (Figure 3-1) shows the years to be nearly identical. This apparent contradiction can be resolved by examining the p_{00} values for the same period; El Niño

years experience a greater value than La Niña years, meaning that while the daily amounts increase during this period for El Niño years, rainfall occurs less frequently.

The relationship between El Niño and La Niña years reverses later in the summer. The La Niña L_1 values are substantially elevated from July to October, while the p_{00} values are lower; this indicates that rainfall is occurring more frequently, and in heavier amounts, during La Niña years in late summer, which is reflected in the total average monthly precipitation graph.

It is interesting to note that, for most of the year, there does not appear to be a great deal of difference in the higher order L-moments (T_2 and T_3 , the coefficient of variation and coefficient of skew, respectively). The most substantial change occurs in March. During this period, the p_{11} , T_2 and T_3 moments are lower in La Niña years than in El Niño years, while the L_1 moment is elevated somewhat. Since the total precipitation amount remains nearly unchanged, it can be concluded that the increase in individual precipitation amounts is countered by a decrease in precipitation frequency. The change in the higher L-moments can be attributed to the upward shift in the L_1 coupled with the decrease in rainfall frequency; fewer data points with a higher mean value would result in lower variability and skewness calculations.

Villa Mills. This station displays a more consistent L-moment seasonal shift than Sanatorio Duran. In this case, the total monthly precipitation graph (Figure 3-2) shows La Niña years experience higher summer amounts, an observation supported by the L_1 graph (Figure 5-9), with La Niña averages being uniformly higher than El Niño years, particularly during the *veranillos*, and in October. During the same period, the T_2 and T_3 (Figure 5-10) values remain fairly consistent for El Niño and La Niña years, as does the

p_{11} parameter. Both the T_2 and T_3 parameters show a slight elevation during the July *veranillos*, which is consistent with the lower overall precipitation during this period; lower daily averages would allow more sporadic higher rainfall amounts to drive up the variability and skew of the underlying distributional form.

Nagatac. The L-moments associated with the Nagatac daily record (Figure 5-11) represent a clear “Pacific” pattern for the region. The winter daily rainfalls are relatively modest and widely spaced, which leads to relatively large readings for T_2 and T_3 (Figure 5-12). This is especially apparent during El Niño years, when it is not uncommon to find completely dry months (e.g., March).

A clear *veranillos* pattern emerges during mid-summer. El Niño year rainfalls are consistently lower than all other year types starting in late June, with a very marked reduction in L_1 values during July. La Niña years, on the other hand, show a substantial increase in daily precipitation values, reaching their maximum at the height of the rainy season in October. These rainfall differences can be linked to the ENSO-based changes in the Trade winds during mid-summer, and the commensurate attenuation of the northward migration of the ITCZ during ENSO events in the late summer.

Time Series Correlation of Transition Probabilities

The results of lag-correlating the p_{00} and p_{11} parameters against the ocean-atmosphere indices discussed in “Methodology” are presented in Figures 5-13 to 5-36. Each figure shows the calculated Pearson correlation coefficient for each month, and lags from -36 to $+36$ months (a negative lag indicates that the index value leads the p_{xx} in time; thus, a lag of -1 would mean that the p_{xx} was correlated with index values which

occurred one month before), along with the degree of statistical significance of the calculated correlation for the same lag-month pairs.

Sanatorio Duran p_{xx} time series correlation. The most notable feature in Figures 5-13 to 5-20 is the overall lack of significant correlation during the summer months for nearly all the indices used. The winter months, especially, December and January, tend to display groups of lags which are more significantly correlated with the p_{xx} . Many of these groupings occur during positive lag, in which the p_{xx} seems to be leading the index conditions by some months.

The SOI tends to be strongly associated with p_{00} and p_{11} only during very narrowly defined periods and lags, notably for December (lag +2 months, +8 months, negative correlation) and January (-16, negative; -1, positive) for p_{00} , and for December (+1, positive; +16, +20, +24, negative) and January (-6, -1, negative) for p_{11} . There is also a p_{11} summer association for June (-16, positive), and for August (-1, negative).

The only significant lag-correlation which occurs with respect to the NAO is a negative association at lag -2 for January. Most other features with reasonably strong significance occur at substantially high lags, and therefore probably represent statistical artifacts (i.e., associations which are related to physical factors other than NAO).

The NINO34 index shows very significant associations with p_{xx} during the winter months, and as with SOI, most of the significant correlation is found during early to mid-winter (November to January), with some strong correlations reaching into February and early March. The p_{00} has a strong positive correlation in December for lags -2 through +12, with the strongest associations appearing in lags +10 through +12; strong negative associations with p_{00} occur at lags -12 to -18. The lag 0 trend tends to carry over into

January through March as advancing lags, culminating with a significant March p_{00} positive correlation at lag -7 months. This corresponds to a negative lag-correlation during the same time period for p_{11} , which seems consistent.

An interesting set of results involves the correlation between p_{00} and p_{11} and the equatorial stratospheric zonal winds (represented by the QBO50 index). The quasi-biennial nature of the QBO50 is clearly evident during the winter/summer transition period in April (for p_{00}), as well as the summer/winter transition period in October and November (for p_{11}). In both cases, periodicities of approximately 27 to 30 months appear in the strong positive and negative correlation pattern for both parameters. For p_{11} , negative correlation exists for October/November for lags of -8 and $+16$ months, and positive correlation exists for the same months for lags of -25 and $+5$ months. For p_{00} , the more significant associations are confined mainly to April; lags of -17 and $+12$ months show the greatest positive results, and -6 and $+24$ months give the strongest negative results. There is also a nearly synchronous (lag -2 months) negative association for p_{00} during the July/August period, which corresponds to the *veranillos*.

Villa Mills p_{xx} time series correlation. When viewed in aggregate, there is a somewhat stronger lag-correlation between the Villa Mills p_{xx} (Figures 5-21 to 5-28) than with Sanatorio Duran for some of the better-understood indices, especially for the NINO34 SST measurements. It is interesting to note, however, that the SOI continues to have few overall significant correlations with the Villa Mills p_{xx} , which suggests that Pacific conditions geographically closer at hand (e.g., NINO34 SST) have a more immediate influence over rainfall frequencies than more distant broad-scale pressure system variability. For the most part, significant SOI positive correlations with p_{11} appear

during the early winter months, and are nearly synchronous, with lags of -3 to $+5$ months. A time of somewhat weaker significance appears during August, approximating the same lags and also positively correlated. The p_{00} shows virtually no period of sustained highly significant correlations.

The NAO (Figures 5-23 and 5-24) shows a similarly weak overall relationship with the p_{xx} measurements at Villa Mills. The few significant correlations (for p_{00} , June, -19 months lag; November, -16 months; for p_{11} , January, $+2$ months) are fairly low (ranging between $\pm .20$ or less. From this it can be deduced that the NAO is of minimal importance to rainfall frequency at Villa Mills, both from an explanatory viewpoint and as a forecasting tool.

NINO34 SSTs show a more coherent signal. The p_{11} shows a stronger overall relationship than p_{00} , with the p_{00} /NINO34 resulting in virtually no correlations with significance exceeding 80% (Figure 5-25). However, the p_{11} correlations (Figure 5-26) show several areas of relatively strong significant correlations: during May (positive correlation, lags $+5$ to $+8$), July and August (negative, lags -4 to $+6$), and December and January (negative, lags -4 to $+14$).

In direct contrast to Sanatorio Duran, Villa Mills exhibits a much weaker relationship with the QBO50 index, both for p_{00} and p_{11} . In addition, the few significant correlations which do appear are for the most part out of phase with those seen at Sanatorio Duran. As might be expected, the strongest relationships for a given month appear to be spaced at an integer multiple of the approximate length of the QBO cycle (27-30 months); the significant February p_{00} negative correlations (Figure 5-27) are at lags of -29 and $+28$ months, thus showing two complete QBO phases. A similar situation

arises for the p_{11} correlations (Figure 5-28) for May (negative, +4 and +32 months; positive, -9 and +22 months; one cycle each) and August/September/October (-30 and +29 months, 2 cycles).

Nagatac p_{xx} time series correlations. Nagatac is located farther to the west, and is therefore more likely to experience a stronger relationship with the indices that describe the state of the Pacific ocean-atmosphere system. The correlations between p_{xx} and the SOI (Figures 5-29 and 5-30) are much more pronounced than at either of the other stations, particularly during the summer months. The near-synchronous negative relationships for p_{00} are most significant starting in July, and extending into the late summer. For each month, lags from -4 to +6 months are most important for p_{00} ; in addition, positive correlations appear for May lags of -14 to -22 months. For mid-to-late summer (August/September), a strong positive correlation between SOI and p_{11} is shown, from 0 to +6 months; the synchronous correlation corresponds to the time frame of the maximum northerly migration of the ITCZ, and can help explain the decrease in rainfall during ENSO events for stations on the Pacific slope. The positive correlation exists for a lag of +5 months demonstrates a relationship between the August p_{11} measurement and the following January SOI, and it is important to note that this is when the SOI reaches its maximum negative value during ENSO events.

The westerly location of Nagatac helps to explain the lack of significant correlations between the p_{xx} and the NAO (Figures 5-31 and 5-32). Such a location would allow Pacific effects to dominate those of the Atlantic basin. While correlations are evident, they tend to be statistically weak, and are therefore of very limited utility for either explanation or forecasting.

The NINO34 correlations (Figures 5-33 and 5-34) provide greater insight into the relationship between p_{xx} and ENSO at Nagatac. Strong positive correlations for p_{00} exist for August (lags -4 to +8) and March/April (lags +6 to +12), and negative p_{00} associations appear in March/April (lags -8 to -2). Strong negative p_{11} correlations exist for March/April (lags +8 to +12 and August/September (centered on lag +4), and positive p_{11} correlations manifest themselves during March/April (lags -6 to -14) and August/September (lags -14 to -24). In each case, the relationships are strongly suggestive of a linkage between the onset of warm water into the NINO34 zone during ENSO events and the migration of the ITCZ during the boreal summer. The correlations during the winter months is probably most closely tied to the effect of ENSO on the *nortes*; however, firm statements are difficult to make due to the scanty nature of the winter rains in western Costa Rica.

There appears to be little in the way of strong relationships between the QBO50 and the p_{xx} for Nagatac (Figures 5-35 and 5-36). As with the other stations, the quasi-biennial nature of the cycle is clearly evident, but the resulting correlations are fairly weak.

Time Series Correlation of L-Moments

Geographic distinctions appear when the L-moments are lag-correlated with the atmospheric indices previously discussed. As a general statement, it is apparent that the amount of daily precipitation occurring on wet days is more strongly correlated to the indices in the eastern part of the transect than in the central and western areas. This is reflected in the L_1 lag-correlation diagrams for the three stations contained in Figures 5-37 through 5-72.

One important finding is common to all three stations: the NAO shows very significant correlations for any of the parameters examined. The conclusion to be drawn, which is discussed in more detail in the next chapter, is that the status of the North Atlantic is of minimal importance in characterizing the region's daily rainfall.

Sanatorio Duran. Sanatorio Duran is more sensitive than the other stations to changes in the conditions in the equatorial Pacific. This seems to be true in both the summer and winter seasons, particularly the late winter. Figure 5-37 shows the near-synchronous positive correlations between the SOI and L_1 from July through September, indicating that during El Niño episodes, the quantity of rainfall falling on wet days is reduced. This seems to correspond to Figure 3-1, which shows that during El Niño periods the total monthly precipitation is reduced. The reverse would be expected as well; La Niña periods would be expected to experience higher intensity daily rain events according to Figure 3-1, and this is supported by Figure 5-37. This trend is supported by the L_1 correlations found for NINO34 (Figure 5-40), in which the synchronous July to September correlations are negative, indicating that warmer Pacific waters are negatively associated with rainfall magnitudes for Sanatorio Duran.

Another feature which appears during the summer months is the association between L_1 and the period preceding it by 12 months (lag -12). During this period, the SOI correlation is negative, while the NINO34 correlations are positive. This is strongly suggestive of an annual "see-saw" effect coming into play, in which the effects of ENSO on the rainfall of Sanatorio Duran reverses itself during the years surrounding the actual event; this supports the findings of previous research which have noted this phenomenon at annual time scales (Waylen et al., 1996).

The QBO (Figure 5-46) has little association with the rainfall magnitudes during the summer months, despite the strong association between p_{11} and the QBO during the late summer months. This peculiar combination of factors seems to point towards the QBO having less impact of storm strength than on storm frequency.

The early-winter correlations show a distinct relationship with the index values preceding the rainfall by six to ten months. This is especially important, because it would indicate that the strength of the *norte* events in eastern Costa Rica is correlated to the occurrence of ENSO events which begin the previous summer. In this case, the correlation is negative, which translates to heavier late-winter rain events during warm-Pacific events initiating the previous summer (i.e., during El Niño years).

Villa Mills. A more subdued pattern emerges for Villa Mills. In the case of L_1 for all of the major indices (Figures 5-49, 5-55, 5-58), there is very little correlation found for any season. This is in contrast to the results for Sanatorio Duran (L_1) and the p_{xx} transition findings for Villa Mills. An important observation, however, is in the behavior of the higher order moments, T_2 and T_3 , which reflect the variability and skew of the rainfall distributions, respectively. For the mid- to late-winter period for both NINO34 (Figure 5-56 and 5-57), there is a strong positive correlation at lag = -10 months, as well as a strong negative correlation for lag = +6 to +8 months. This corresponds well to the observed behavior of the *nortes* during and following an El Niño event; the +8 month negative lag shows that higher rainfall values are depressed during the year of an ENSO event, whereas the year following (-10 month lag) the more intense rainfalls resume. In addition, the NAO shows extremely few high-significance correlations, which parallels the finding at the other stations.

The QBO50- L_1 relationship is more difficult to interpret. The only clear interpretation that can be made is that the 27-30 month cycle of the QBO is clearly evident in the L_1 , T_2 , and T_3 graphs (Figure 5-58 to 5-60). The correlations are strongest for the T_2 and T_3 moments; early winter positive associations with the QBO result with a lag of -16 to -12 months, during early summer (May) for near synchronous correlations, and during the *veranillos* (July) period for lags of -9 to -12 months.

Nagatac. The Nagatac L_1 correlations (Figure 5-61) show few similarities with the p_{00} and p_{11} correlations. The three lag-regions which show significant correlations are March (negative, lag 0 to +10), July (positive, lag +1), and October (negative, lag -8 to -12). These same regions appear as highly correlated in the NINO34 with the opposite associations; the March region is strongly positive, July is strongly negative, and October is positive. None of these areas are significant on the p_{xx} lag-correlations, which seems to suggest that changes in the western Pacific pressure dipole, and associated shifts in warm water across the equatorial Pacific, manifests themselves in different ways for different months. The relatively stronger significance measurements in the NINO34 results seems to indicate that the region responds more readily to alterations in the Pacific SST's rather than with changes in the pressures over the central and western Pacific.

The QBO50- L_1 correlation results (Figure 5-67) show a much stronger association between the zonal stratospheric winds and the intensity of rainfall in western Costa Rica. The quasi-biennial oscillation is clearly evident in the spring months (March/April), and even more strongly from October to December. This strongly suggests that a physical association is present; unfortunately, this area of research is in a developing state, and few conclusive studies have been made concerning the linkage between climate

anomalies and stratospheric wind flow. Nevertheless, the associations present in the record indicate that QBO50 could represent an excellent forecasting variable even if a physical linkage cannot be found.

Model Choice, Validation, and Rainfall Simulation

Initial Model Choice. The first step in validating a particular statistical intensity model is to identify the candidate distributions to be tested. As stated in the “Methodology” chapter, 11 candidate distributions were identified, which range in complexity from a single parameter exponential model to a more complex five-parameter Wakeby distribution. Two months (January, July) were chosen for each station, and year-types segregated into “ENSO” and “non-ENSO” years based upon the Kiladis and Diaz classification used for the earlier tests. The results of the Chi-squared initial elimination process are shown in Table 5-4.

The results clearly show a bias towards the higher-parameter distributions. The exception is the choice of the gamma distribution, which exceeded both the Generalized Logistic and Generalized Extreme Value distributions in choice likelihood based on the Chi-squared tests. It is interesting to note that the Gamma distribution is the only two-parameter distribution to include scale and shape as the parameters, rather than the location and scale parameters for the Normal and Gumbel distributions.

Table 5-4. Tallied results of distributional Chi-squared tests.

distribution (# of params.)	January ENSO years	January non-ENSO years	July ENSO years	July non-ENSO years	TOTAL COUNT
Exponential (1)	1	0	1	0	2
Gamma (2)	3	0	3	3	9 *
Gumbel (2)	0	0	0	0	0
Normal (2)	0	0	0	0	0
Generalized Extreme Value (3)	3	2	1	1	8
Generalized Logistic (3)	4	2	0	1	7
Generalized Normal (3)	2	3	2	3	10 *
Generalized Pareto (3)	3	3	3	3	12 *
Pearson III (3)	2	2	3	3	10 *
Kappa (4)	3	3	3	3	12 *
Wakeby (5)	2	3	2	2	9 *

* indicates that the distribution was selected for "final six" consideration

Rainfall simulations. The next step in deciding upon a proper distributional form is to use the calculated parameters (both intensity and frequency) and perform simulations based upon the candidate distributions and the Markov model parameters. The simulations are run for each station/distribution/year-type/month combination, with

1000 iterations being used to calculate the total monthly precipitation. The results are plotted in Figures 5-73 to 5-78.

It is apparent that most of the distributions give a very close correspondence to the actual average monthly totals for all stations. Sanatorio Duran and Villa Mills show the greatest overall correspondence, with Nagatac monthly totals being the most difficult to replicate. This is especially true during the *veranillos* period of El Niño years, and for the latter half of summer during La Niña years. It is noteworthy that, in general, all the distributions seem to underestimate the actual monthly totals during the summer at Nagatac; it is likely that this reflects the very high totals that routinely fall in the western part of Costa Rica during this period. The underestimation during El Niño years is most significant during the *veranillos*, when the estimated values are about half of the actual amounts; this is probably attributable to the sudden decrease in daily rainfall amounts during the onset of the increased Trade winds, making predictions at the monthly scale highly variable.

As mentioned, nearly all of the distributions do an adequate job of modeling the precipitation at the monthly scale. The notable exception is the Kappa distribution, which radically underestimates amounts for both Villa Mills and Nagatac during the spring months (March/April). This is sufficient evidence to lead to the conclusion that the Kappa distribution is not a preferred model for daily precipitation, and can be removed from further consideration.

Extreme value analysis. Since the monthly rainfall simulation failed to distinguish among the candidate distributions, it is necessary to consider the effect that each of the models has on simulating the extreme daily precipitation events. The “bootstrap”

technique outlined by Wilks (1993) is employed, as described in the Methodology chapter. The data are segregated as in the Chi-squared test, and 500 bootstrap simulations are performed. The distribution of choice will be the one which provides the best fit between the median value and actual value (“median-to-actual fit”), as well as the least sampling variability (i.e., the “tightest” fit). The results are shown in Figures 5-79 to 5-84. In each of the bootstrap diagrams, the five largest daily rainfall events are shown on each distribution graph as asterisks, along with a box-whisker plot showing the 5%, 25%, 75%, and 95% boundaries for the 500 sample calculations representing the rainfall estimate for that value.

Several notable observations can be seen in the results. Overall, it is apparent that the distributions with the largest number of parameters do not provide the best fit to the extreme values in any of the cases. In fact, in many cases, the best fit is accomplished by the distribution with the fewest parameters (Gamma) gives the best overall fit in terms of median-to-actual fit, as well as minimizing the sampling variability. In addition, both the Gamma and Pearson III models seem to mirror the others responses; this is probably due to the fact that the Pearson III distribution is nearly identical to the Gamma with the addition of a location parameter.

Sanatorio Duran shows a tendency towards consistent overestimation as well, but the departures are minimized by using either a Gamma or Pearson III distribution. (In fact, the Gamma shows a slight tendency to underestimate, rather than overestimate, the largest values.) The results are similar for both January and July, and for El Niño and non-El Niño months. In addition, the Gamma and Pearson III models show the lowest

amount of sampling variation. For all cases, it is clear that the Generalized Logistic, Generalized Normal, Kappa, and Wakeby distributions can be discounted.

Villa Mills during July experiences at least one very large rainfall in both ENSO and non-ENSO years, and the resulting distributional fits are quite imprecise. The January estimates are much more consistent, in that it is possible to find models with a reasonable degree of fit to the data. In this case, the Pearson III model provides the highest degree of correspondence to the actual amounts as well as the second-lowest sampling variability (behind the Gamma distribution). The Pearson III and Gamma distributions also seem to give an adequate fit to the extreme values when the largest value is disregarded.

The Nagatac extreme values in July have a very difficult time being fit by any of the model choices, despite the relative consistency of the actual amounts. The degree of deviation from the actual amounts is greatest for ENSO years, when there is a clear tendency to overestimate the largest values. The degree to which the deviation occurs is minimized with the Gamma and Pearson III models, both in terms of the actual-to-median difference and among the sampling variability of the bootstrap samples. The Kappa model shows the greatest amount of sampling variation, especially during January, when the rain events seldom occur.

Parameter Forecasting

The one-year forecast results based upon the creation of individual month-by-month regression models for each parameter, as well, as the ARIMA forecast results, are shown in Tables 5-1 to 5-9, and graphically illustrated in Figures 5-85 to 5-87, and the ARIMA results (with 95% confidence intervals illustrated) are shown alone in Figures 5-

88 to 5-90. The results show that, in general, the ARIMA model performs better in tracking the actual parameter values. This is especially evident when the month-to-month transitions are relatively “smooth,” although this is not exclusively the case; it appears that, because the multiple regression models are individually created for each month without regard for the predictors used in adjacent months, large inter-monthly variations can occur.

One of the most striking observations concerning the multiple regression models is the frequency with which the NAO index is used as a predictor. This is in stark contrast to the individual monthly time-series correlations discussed earlier, where it was found that the NAO had very low significant correlations with the parameters. From this result, it may be inferred that the strong showing of the NAO in the prediction models indicates there is a significant amount of interaction (both statistical and physical) between the NAO and the processes reflected in the SOI, NINO34, and QBO indices.

A further observation about the multiple regression forecast models concerns the apparent lack of a relationship between the multiple-R and adjusted-R² measurements and the ability of the models to accurately forecast the parameters. The R and adjusted-R² values seem to be more related to the number of model predictors, and a strong R appears to give no guarantee that the model will perform adequately as a forecast tool. In fact, it was often the case that the models gave impossible values as forecasts (i.e., transition probabilities greater than one or less than zero, or L₁ values less than zero). The ARIMA model showed no such tendency, indicating it is a more “stable” prediction scheme.

Table 5-1. Forecast model results for Sanatorio Duran L₁.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model*	R	adjusted R ²	predicted (mult. reg.)		
JAN	-51.6431 + 2.017 NIN ₈ + 0.745 NAO ₁₈ - 0.457 NAO ₇ - 0.307 NAO ₁₂	0.892	0.753	3.449	4.324	4.254
FEB	2.645 - 0.475 NAO ₁₈ + 0.772 SOI ₂₀	0.749	0.521	1.843	2.511	0.780
MAR	2.653 - 0.719 NAO ₁₇ - 0.816 SOI ₆	0.694	0.481	2.514	1.853	0.400
APR	2.603 - 0.129 QBO ₉ + 2.113 SOI ₂₆ - 2.914 SOI ₂₄	0.754	0.506	-0.899	2.335	1.565
MAY	9.373 + 2.530 SOI ₂₂ - 1.559 SOI ₁₆ + 2.424 SOI ₁₁ + 0.776 NAO ₂₄	0.755	0.489	12.096	9.385	7.879
JUN	5.597 +1.627 NAO ₃₅ - 2.649 SOI ₁₅ - 0.224 QBO ₂₀ + 0.190 QBO ₁₆ - 0.075 QBO ₃₆ + 0.465 NAO ₂₅	0.927	0.814	8.303	5.974	8.271
JUL	-35.666 + 1.447 NAO ₁₀ - 0.063 QBO ₃₃ + 3.911 NIN ₉ + 0.462 NAO ₂₂ - 2.314 NIN ₆ + 0.535 NAO ₃₆	0.912	0.778	6.115	6.065	5.308
AUG	9.328 - 2.077 SOI ₁₂ + 1.208 NAO ₉ + 0.730 NAO ₁₁	0.679	0.387	9.074	6.984	8.069
SEP	11.063 + 1.030 NAO ₁₂	0.436	0.159	10.650	9.357	5.400
OCT	11.722 - 2.678 SOI ₁₉ + 0.852 NAO ₁₇ + 0.919 NAO ₂₇ + 0.775 NAO ₁₁ - 0.056 QBO ₂₃	0.815	0.584	14.311	10.775	8.500
NOV	10.505 - 0.105 QBO ₈ - 0.525 NAO ₈ - 0.677 NAO ₃₀ + 1.115 SOI ₃₃ - 2.492 SOI ₃₁ - 1.927 SOI ₇ - 0.063 QBO ₃₅ - 0.511 NAO ₁₂ + 0.082 QBO ₇	0.947	0.841	9.180	8.500	5.700
DEC	-84.567 + 3.296 NIN ₇ - 0.048 QBO ₉ + 1.234 SOI ₇	0.755	0.514	6.603	6.025	3.810

* The number of months of lag applied to atmospheric indices as a model predictor is indicated by the number following index (e.g., “NAO₁₀” indicated the NAO lagged by 10 months)

Table 5-2. Forecast model results for Sanatorio Duran p₀₀.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	0.791 + 0.054 SOI ₁₈ - 0.002 QBO ₁₃	0.708	0.458	0.795	0.805	0.709
FEB	0.849 + 0.070 SOI ₁₆ + 0.034 NAO ₂₄ - 0.035 NAO ₆ - 0.077 SOI ₂₃ + 0.044 SOI ₂₇ - 0.010 NAO ₃₀ - 0.001 QBO ₁₄ + 0.007 NAO ₃₃	0.962	0.887	0.958	0.853	0.788
MAR	3.055 - 0.028 NAO ₈ + 0.029 NAO ₁₄ + 0.034 NAO ₃₃ - 0.081 NIN ₃₂ - 0.027 SOI ₂₅	0.851	0.653	1.049	0.837	0.880
APR	0.818 + 0.089 SOI ₂₁ + 0.043 NAO ₁₄ + 0.036 NAO ₁₈	0.804	0.595	1.056	0.799	0.417
MAY	0.531 + 0.133 SOI ₃₄ - 0.067 NAO ₂₂ + 0.093 SOI ₁₀ - 0.003 QBO ₈	0.849	0.667	0.692	0.360	0.000
JUN	0.502 + 0.007 QBO ₁₂ - 0.142 SOI ₂₆ + 0.067 NAO ₈ + 0.058 NAO ₃₅ + 0.057 SOI ₃₅	0.879	0.716	0.605	0.569	0.462
JUL	3.115 + 0.086 NAO ₂₇ + 0.080 NAO ₁₉ + 0.108 SOI ₇ - 0.103 NIN ₂₁ - 0.082 SOI ₁₆	0.911	0.787	0.582	0.293	0.286
AUG	0.453 0.152 NAO ₃₅ - 0.070 NAO ₂₇ + 0.041 NAO ₃₂ - 0.056 SOI ₂₀ - 0.006 QBO ₂₁ - 0.055 NAO ₂₈ + 0.006 QBO ₁₆ - 0.111 SOI ₂₄	0.940	0.829	0.667	0.274	0.400
SEP	0.345 + 0.087 NAO ₂₁ + 0.077 NAO ₃₃ + 0.055 NAO ₃₆ + 0.055 NAO ₂₃ + 0.036 NAO ₃₅	0.838	0.634	0.139	0.321	0.600
OCT	0.339 - 0.215 SOI ₂₃ + 0.085 SOI ₂₀	0.746	0.520	0.221	0.217	0.200
NOV	0.579 + 0.086 SOI ₁₆ - 0.061 NAO ₃₂ - 0.099 SOI ₁₉ + 0.027 NAO ₂₂	0.763	0.506	0.512	0.467	0.481
DEC	0.700 + 0.049 SOI ₁₇	0.508	0.229	0.725	0.689	0.571

Table 5-3. Forecast model results for Sanatorio Duran p₁₁.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	-2.306 + 0.005 QBO ₆ + 0.050 NAO ₁₈ - 0.052 NAO ₃₆ + 0.109 NIN ₁₆ + 0.037 NAO ₃₁ + 0.026 NAO ₃₅ + 0.016 NAO ₈	0.954	0.876	0.477	0.645	0.543
FEB	2.364 - 0.041 NAO ₃₀ - 0.038 NAO ₉ - 0.029 NAO ₁₃ - 0.072 NIN ₇	0.810	0.588	0.326	0.575	0.458
MAR	0.213 - 0.043 SOI ₁₅ + 0.108 NAO ₂₁ + 0.091 NAO ₁₁ - 0.065 NAO ₃₀	0.849	0.647	0.361	0.379	0.600
APR	0.367 - 0.072 NAO ₁₅ + 0.007 QBO ₂₈ + 0.072 NAO ₃₄ + 0.005 QBO ₁₈	0.750	0.475	0.165	0.395	0.529
MAY	2.235 - 0.041 NAO ₂₈ + 0.038 NAO ₃₃ + 0.040 NAO ₉ - 0.056 NIN ₂₆	0.820	0.610	0.787	0.683	0.893
JUN	0.730 - 0.022 NAO ₆ + 0.055 SOI ₁₅ - 0.023 NAO ₃₄ - 0.022 NAO ₂₁	0.739	0.459	0.610	0.833	0.625
JUL	0.642 - 0.143 SOI ₃₆ + 0.143 SOI ₃₄ - 0.084 NAO ₁₉ + 0.057 NAO ₁₅ + 0.041 NAO ₂₄ + 0.030 NAO ₁₇	0.859	0.655	1.330	0.774	0.783
AUG	0.764 + 0.005 QBO ₂₅ - 0.058 NAO ₂₅	0.659	0.385	0.804	0.756	0.880
SEP	0.724 - 0.040 NAO ₃₂ - 0.073 NAO ₃₆ + 0.003 QBO ₃₄ + 0.057 NAO ₁₆ + 0.103 SOI ₂₇ - 0.035 SOI ₂₀ + 0.027 NAO ₃₀	0.888	0.715	0.610	0.768	0.958
OCT	3.937 + 0.008 QBO ₂₀ + 0.052 NAO ₃₃ - 0.026 NAO ₁₂ - 0.116 SOI ₆ - 0.117 NIN ₇ - 0.021 NAO ₂₁ - 0.003 QBO ₆ - 0.002 QBO ₂₂	0.929	0.802	0.500	0.826	0.800
NOV	0.954 + 0.005 QBO ₂₆ - 0.027 NAO ₁₃ + 0.019 NAO ₁₆ + 0.039 NAO ₃₃ - 0.015 NAO ₂₂ + 0.052 NIN ₁₈ - 0.064 SOI ₂₅ - 0.057 NIN ₃₀ + 0.036 SOI ₂₉	0.951	0.853	0.861	0.747	0.778
DEC	0.638 + 0.628 NAO ₃₃ + 0.033 QBO ₂₇ + 0.028 NAO ₆ + 0.086 SOI ₆ - 0.032 NAO ₂₄ - 0.042 SOI ₂₉ + 0.024 SOI ₃₃	0.892	0.720	0.560	0.672	0.769

Table 5-4. Forecast model results for Villa Mills L₁.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	4.460 - 1.597 SOI ₁₂ - 0.245 QBO ₂₃ - 0.059 QBO ₃₂ + 0.168 QBO ₂₄	0.793	0.564	2.610	2.410	5.541
FEB	-43.858 + 1.707 NIN ₉ + 0.531 SOI ₁₁	0.583	0.289	3.440	1.395	1.858
MAR	2.992 + 0.520 SOI ₂₅	0.416	0.140	3.564	1.617	1.064
APR	-211.876 + 13.418 NIN ₁₁ + 2.122 NAO ₃₄ + 1.918 NAO ₂₅ + 0.110 QBO ₂₈ - 5.472 NIN ₁₀ - 0.740 NAO ₁₆	0.927	0.818	18.476	5.765	5.000
MAY	12.049 + 1.190 NAO ₁₇ + 1.162 SOI ₁₄ - 1.545 NAO ₂₅ - 1.727 SOI ₁₆ + 1.077 SOI ₇ + 0.824 NAO ₃₂ + 0.543 NAO ₂₉ + 0.436 NAO ₁₁	0.892	0.718	7.236	10.237	11.842
JUN	11.954 - 2.299 SOI ₁₃ + 1.148 NAO ₁₀ - 0.082 QBO ₃₃ + 0.904 NAO ₁₈	0.855	0.689	5.980	10.495	15.124
JUL	6.975 - 0.137 QBO ₁₁ + 1.146 NAO ₁₃ - 0.118 QBO ₃₂ + 1.157 NAO ₁₁ + 0.934 SOI ₃₀ - 1.433 SOI ₁₅ + 0.612 NAO ₁₂ + 0.038 QBO ₂₅ - 0.485 NAO ₂₅	0.949	0.854	4.176	8.201	5.292
AUG	-77.284 - 4.894 SOI ₁₆ + 2.313 NAO ₁₈ + 1.728 SOI ₂₉ - 1.914 SOI ₇ + 5.215 SOI ₁₂ - 1.077 NAO ₂₇ - 1.019 NAO ₂₆ + 3.502 NIN ₁₁ - 0.505 NAO ₈ + 0.031 QBO ₂₈	0.974	0.921	20.781	12.323	13.075
SEP	13.316 - 1.617 NAO ₃₆ - 0.121 QBO ₁₄ - 2.089 SOI ₁₇ - 0.084 QBO ₃₀	0.869	0.717	8.597	12.948	14.258
OCT	208.310 - 1.780 SOI ₂₉ - 5.784 SOI ₁₃ - 7.249 NIN ₁₄ - 0.138 QBO ₁₄ + 1.051 NAO ₂₆ + 0.875 NAO ₉ + 0.859 NAO ₂₃ - 1.698 SOI ₆	0.946	0.850	16.491	14.430	12.748
NOV	9.036 - 0.114 QBO ₁₃ - 0.092 QBO ₂₆ - 0.561 NAO ₆ - 0.658 NAO ₃₁ - 0.686 SOI ₂₀	0.864	0.691	8.148	10.509	7.507
DEC	5.334 - 0.112 QBO ₉ - 0.154 QBO ₂₂ + 0.711 NAO ₃₃ - 0.075 QBO ₂₉ + 0.491 NAO ₈ + 0.469 SOI ₂₆ - 0.481 NAO ₁₇ + 0.115 QBO ₂₃	0.927	0.803	8.755	6.580	3.508

Table 5-5. Forecast model results for Villa Mills p00.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	1.952 + 0.032 NAO ₂₆ - 0.046 NIN ₁₆ - 0.025 NAO ₁₉	0.722	0.462	0.692	0.699	0.731
FEB	1.727 - 0.041 NAO ₂₀ + 0.026 NAO ₂₇ - 0.002 QBO ₇ - 0.038 NIN ₁₂	0.844	0.665	0.682	0.673	0.724
MAR	2.158 - 0.049 NAO ₂₁ - 0.003 QBO ₈ - 0.032 NAO ₁₇ - 0.053 NIN ₃₁	0.818	0.611	0.807	0.589	0.750
APR	0.684 + 0.048 NAO ₃₂ + 0.003 QBO ₁₆ + 0.081 SOI ₂₇ - 0.074 SOI ₂₉	0.800	0.576	0.820	0.564	0.625
MAY	0.263 + 0.109 SOI ₃₄ + 0.014 QBO ₁₈ - 0.011 QBO ₂₀	0.685	0.403	0.039	0.141	0.000
JUN	0.176 - 0.057 NAO ₂₇ + 0.196 SOI ₂₄ - 0.121 SOI ₂₅	0.821	0.633	-0.041	0.058	0.800
JUL	0.022 - 0.006 QBO ₉ - 0.006 QBO ₂₂ - 0.037 NAO ₇ - 0.005 QBO ₃₄ + 0.024 NAO ₁₈ - 0.040 NAO ₂₃ - 0.029 NAO ₁₃	0.860	0.653	-0.200	0.110	0.000
AUG	0.330 - 0.047 NAO ₁₅	0.389	0.117	0.414	0.164	0.143
SEP	0.124 + 0.077 SOI ₇ + 0.009 QBO ₂₃ - 0.008 QBO ₂₇ - 0.050 NAO ₂₂ + 0.007 QBO ₃₄ + 0.074 SOI ₃₀ + 0.061 NAO ₃₃ + 0.029 NAO ₂₅ + 0.061 NAO ₃₃ + 0.029 NAO ₂₅ + 0.032 NAO ₁₇ - 0.037 SOI ₉	0.982	0.930	0.391	0.170	0.000
OCT	-1.890 - 0.052 NAO ₂₈ + 0.035 NAO ₃₄ - 0.026 NAO ₇ - 0.073 NIN ₃₁	0.875	0.706	0.078	0.097	0.000
NOV	-2.406 + 0.054 NAO ₁₅ + 0.024 NAO ₂₉ + 0.115 NIN ₂₂ - 0.065 NAO ₁₆ + 0.006 QBO ₂₃ - 0.008 QBO ₂₁ + 0.037 NAO ₁₂ + 0.002 QBO ₁₆ - 0.053 SOI ₃₅ + 0.039 NAO ₂₈ + 0.031 SOI ₃₃ + 0.033 SOI ₂₀ + 0.002 QBO ₂₆	0.988	0.956	1.038	0.541	0.478
DEC	4.055 - 0.058 NAO ₁₉ - 0.050 NAO ₁₇ - 0.125 NIN ₉ - 0.028 NAO ₁₈ + 0.021 NAO ₃₁	0.860	0.683	0.705	0.666	0.545

Table 5-6. Forecast model results for Villa Mills p₁₁.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg)		
JAN	0.579 + 0.014 QBO ₃₂ + 0.027 NAO ₁₉ - 0.029 NAO ₂₅ - 0.027 NAO ₂₈ - 0.024 QBO ₃₁ + 0.0188 QBO ₃₀ - 0.002 QBO ₈ + 0.016 NAO ₁₁ + 0.009 NAO ₂₃ - 0.012 NAO ₂₉ - 0.013 NAO ₃₃	0.984	0.947	0.594	0.726	0.632
FEB	-1.796 + 0.065 NAO ₇ - 0.048 NAO ₃₆ + 0.018 NAO ₂₃ + 0.115 SOI ₁₇ + 0.087 NIN ₁₆	0.852	0.667	0.625	0.638	0.531
MAR	0.474 + 0.066 NAO ₃₅	0.354	0.091	0.315	0.617	0.500
APR	0.590 + 0.041 NAO ₃₅ - 0.129 SOI ₂₁ + 0.110 SOI ₂₃ - 0.033 NAO ₉	0.689	0.383	0.629	0.668	0.538
MAY	4.266 - 0.125 NIN ₁₄ - 0.026 NAO ₂₁ + 0.044 SOI ₃₆	0.763	0.534	0.797	0.872	1.000
JUN	0.011 - 0.027 NAO ₃₃ + 0.034 NIN ₉ - 0.013 NAO ₆ - 0.011 NAO ₁₈ + 0.001 QBO ₆	0.814	0.592	0.882	0.970	0.958
JUL	2.548 + 0.046 NAO ₁₂ - 0.008 SOI ₃₃ + 0.004 QBO ₃₆ + 0.002 QBO ₂₃ - 0.060 NIN ₁₇ - 0.016 NAO ₆	0.874	0.700	0.983	0.890	0.800
AUG	0.913 + 0.003 QBO ₃₆ + 0.028 NAO ₁₄ - 0.039 SOI ₃₄ - 0.001 QBO ₈	0.737	0.463	0.980	0.810	0.739
SEP	-0.829 - 0.002 QBO ₃₃ + 0.064 NIN ₂₈ + 0.030 SOI ₁₃ - 0.022 SOI ₇ - 0.001 QBO ₂₁	0.883	0.734	0.821	0.888	0.840
OCT	0.939 - 0.027 SOI ₃₆	0.585	0.317	0.901	0.948	0.929
NOV	0.864 - 0.014 NAO ₃₀ + 0.019 NAO ₁₇ + 0.002 QBO ₃₀	0.645	0.345	0.903	0.886	0.806
DEC	0.772 + 0.004 QBO ₃₁ + 0.024 NAO ₁₈ - 0.030 NAO ₂₄ + 0.028 NAO ₁₁ - 0.002 QBO ₇ - 0.012 NAO ₂₃	0.869	0.689	0.911	0.806	0.711

Table 5-7. Forecast model results for Nagatac L₁.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	-49.860 - 2.255 NAO ₁₇ - 0.233 QBO ₂₅ + 2.167 NIN ₈ + 0.150 QBO ₂₉ + 0.093 QBO ₁₈	0.948	0.867	6.390	4.203	2.743
FEB	11.737 + 0.316 QBO ₉ - 2.400 NAO ₆ - 3.080 NAO ₁₄ - 1.270 NAO ₃₃ - 1.641 NAO ₉ - 0.727 NAO ₂₅	0.993	0.976	3.190	4.441	2.102
MAR	21.300 + 0.464 QBO ₇ + 3.301 NAO ₈ + 0.009 QBO ₃₀	0.999	0.999	-1.740	5.640	5.300
APR	10.811 - 0.814 QBO ₁₂ - 4.498 NAO ₂₂ + 5.095 SOI ₁₃ + 3.818 NAO ₃₂ - 3.269 NAO ₃₄ - 2.105 NAO ₂₄ + 0.155 QBO ₁₈ + 0.976 NAO ₁₃ - 0.058 QBO ₂₄ + 1.194 SOI ₂₃ - 0.233 SOI ₂₇ - 0.005 QBO ₁₀ - 0.007 NIN ₃₄	0.999	0.999	15.400	6.583	8.112
MAY	18.550 + 2.340 NAO ₁₁ - 5.822 SOI ₂₄ - 2.230 NAO ₂₁ - 2.240 NAO ₁₃	0.813	0.589	23.620	12.862	16.375
JUN	16.232 + 2.472 NAO ₁₃ + 2.149 NAO ₂₄	0.700	0.442	13.110	11.199	29.308
JUL	11.950 - 2.033 NAO ₃₁	0.468	0.183	9.710	9.300	14.150
AUG	16.590 - 7.095 SOI ₁₃ + 8.742 SOI ₁₄ - 1.447 NAO ₂₈ - 2.053 SOI ₂₈ + 1.572 NAO ₃₁ + 1.627 NAO ₂₇ + 1.393 NAO ₂₃ - 1.796 SOI ₁₈ + 2.069 SOI ₂₆ - 0.664 SOI ₃₁	0.985	0.945	14.482	11.512	30.885
SEP	13.669 - 4.226 NAO ₁₃ + 2.041 NAO ₃₁ - 0.331 QBO ₁₃ + 10.485 SOI ₁₆ - 0.237 QBO ₂₈ - 2.798 SOI ₆ + 1.602 NAO ₇	0.944	0.838	16.522	9.324	23.447
OCT	-307.386 + 12.282 NIN ₁₅ + 7.080 SOI ₁₅	0.673	0.396	15.300	18.807	40.121
NOV	15.597 + 2.528 NAO ₃₆ - 0.130 QBO ₃₅ - 1.364 NAO ₃₂ - 3.960 SOI ₁₇	0.806	0.577	23.601	14.115	27.433
DEC	91.200 + 3.634 NIN ₃₄ + 1.698 NAO ₂₈ - 6.548 NIN ₇ - 1.941 SOI ₉	0.807	0.575	8.027	9.270	11.770

Table 5-8. Forecast model results for Nagatac p00.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	0.943 -0.015 NAO21 + 0.009 NAO24 - 0.009 NAO19	0.720	0.456	0.949	0.899	0.900
FEB	0.931 + 0.002 QBO35 + 0.015 NAO30 + 0.019 NAO31 - 0.001 QBO21 + 0.008 NAO27 - 0.016 SOI14 - 0.001 QBO33 - 0.001 QBO8 - 0.007 SOI35	0.906	0.720	1.011	0.960	0.854
MAR	0.944 - 0.005 QBO20 + 0.017 NAO7 + 0.015 NAO29 + 0.004 QBO19	0.761	0.499	0.860	0.975	0.750
APR	0.901 + 0.040 NAO16 + 0.027 NAO6 + 0.022 NAO20	0.748	0.499	0.960	0.956	0.810
MAY	4.133 + 0.173 SOI24 - 0.022 QBO28 + 0.015 QBO30 - 0.130 NIN33 - 0.055 NAO13	0.927	0.824	-0.030	0.648	0.000
JUN	0.204 - 0.005 QBO20 + 0.058 NAO28 - 0.068 NAO33	0.742	0.489	0.077	0.521	0.333
JUL	0.448 + 0.066 NAO18 + 0.075 NAO34 + 0.081 NAO26 + 0.054 NAO27 - 0.058 NAO12 - 0.041 NAO14	0.911	0.733	0.427	0.628	0.500
AUG	0.490 + 0.007 QBO13 + 0.051 NAO6	0.680	0.411	0.210	0.455	0.500
SEP	0.168 + 0.127 NAO24 - 0.067 NAO27	0.730	0.471	-0.090	0.532	0.000
OCT	0.122 - 0.007 QBO22 - 0.172 SOI17	0.722	0.468	0.230	0.309	0.000
NOV	0.726 + 0.196 SOI36 - 0.095 SOI30 + 0.038 NAO12 - 0.028 NAO36 + 0.054 NIN36	0.934	0.840	0.502	0.744	0.854
DEC	0.872 + 0.021 NAO23 + 0.028 SOI9	0.681	0.419	0.908	0.859	0.915

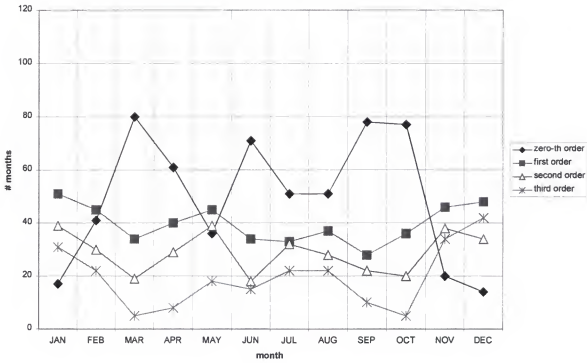
Table 5-9. Forecast model results for Nagatac p₁₁.

Month	Multiple Regression				predicted (ARIMA)	actual value
	Model	R	adjusted R ²	predicted (mult. reg.)		
JAN	10.294 - 0.011 QBO ₉ - 0.011 QBO ₂₇ + 0.015 NAO ₃₂ + 0.459 NIN ₃₀ + 0.220 SOI ₂₈ + 0.007 QBO ₂₈ - 0.072 NIN ₁₄ + 0.027 NAO ₂₁	0.968	0.904	0.496	0.062	0.286
FEB	0.343 - 0.080 NAO ₁₁ - 0.065 NAO ₃₂ - 0.056 NAO ₈ - 0.053 NAO ₁₉ - 0.041 NAO ₁₂ - 0.003 QBO ₃₆	0.976	0.926	-0.093	0.169	0.000
MAR	0.389 - 0.189 NAO ₁₀ + 0.199 SOI ₃₀ + 0.107 NAO ₉ - 0.001 QBO ₁₁ + 0.001 SOI ₁₆	0.999	0.999	0.824	0.325	0.000
APR	0.396 + 0.096 NAO ₂₇ - 0.073 NAO ₁₃	0.771	0.536	0.298	0.482	0.500
MAY	0.542 + 0.073 NAO ₃₃ + 0.044 NAO ₁₁ - 0.048 NAO ₇ + 0.126 SOI ₂₁ + 0.005 QBO ₃₁ - 0.029 NAO ₂₈	0.901	0.754	0.550	0.626	0.667
JUN	-0.509 - 0.038 NAO ₂₈ + 0.070 SOI ₂₅ + 0.026 NAO ₂₀ + 0.048 NIN ₂₄ - 0.014 NAO ₂₉ + 0.014 NAO ₁₀	0.888	0.723	0.750	0.682	0.826
JUL	-5.171 + 0.216 NIN ₃₆ - 0.078 SOI ₁₇ + 0.068 SOI ₃₀ - 0.066 SOI ₂₈ + 0.063 SOI ₂₁	0.899	0.761	0.565	0.689	0.833
AUG	0.687 - 0.128 SOI ₁₆ + 0.057 NAO ₁₄ - 0.039 NAO ₁₈ + 0.044 NAO ₂₃ + 0.020 NAO ₁₀ + 0.016 NAO ₁₁	0.963	0.903	0.968	0.730	0.885
SEP	0.824 - 0.076 SOI ₁₉ + 0.032 NAO ₁₂ + 0.040 NAO ₁₅ + 0.025 NAO ₁₆ + 0.051 SOI ₂₂	0.912	0.785	0.898	0.662	1.000
OCT	2.046 + 0.076 SOI ₂₆ - 0.045 NIN ₃₆ + 0.026 NAO ₃₀	0.776	0.543	0.785	0.648	0.893
NOV	7.182 - 0.040 NAO ₃₄ - 0.141 NIN ₆ - 0.099 NIN ₂₆	0.733	0.475	0.558	0.553	0.810
DEC	0.467 + 0.074 NAO ₃₀ - 0.051 NAO ₃₅	0.675	0.406	0.343	0.451	0.650

Results Overview

The results indicate that the daily rainfall intensity and occurrence characteristics of the region are quite complex, both from a temporal and spatial standpoint. The simulation results show that, despite the inherent complexity of the processes, the behavior can be modeled, and that many of the mechanisms that underlie the system can be adequately modeled using the candidate distributions described. Attempts to apply these concepts to forecasting using multiple regression and ARIMA methods results in considerable divergence from the withheld data, which reiterates the underlying complexity of the system.

AIC -- All Year Types



BIC -- All Year Types

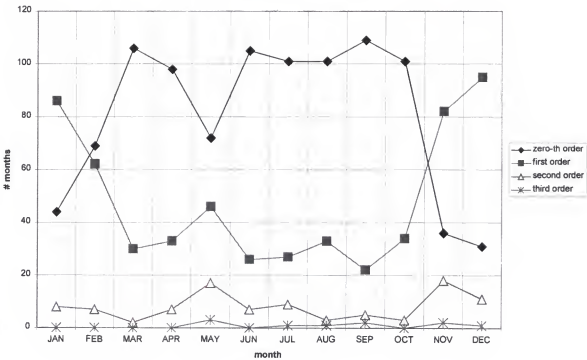
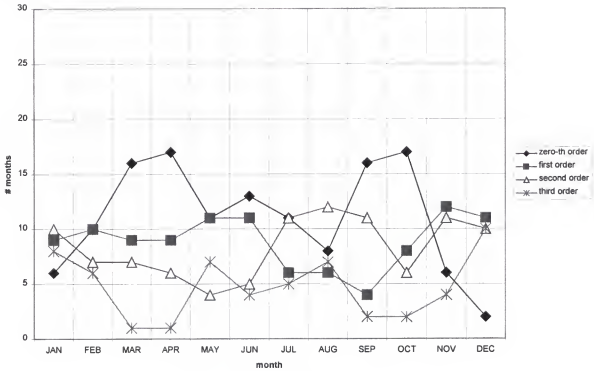


Figure 5-1. Monthly AIC and BIC statistics, all years

AIC – El Niño Years



BIC – El Niño Years

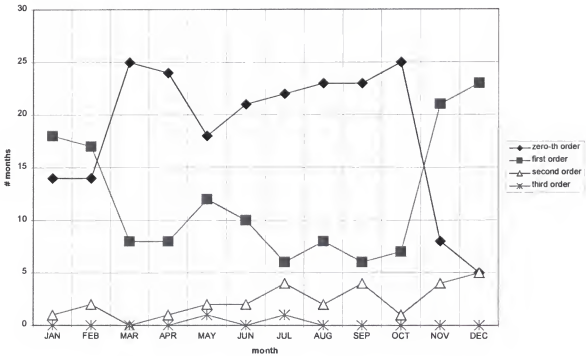
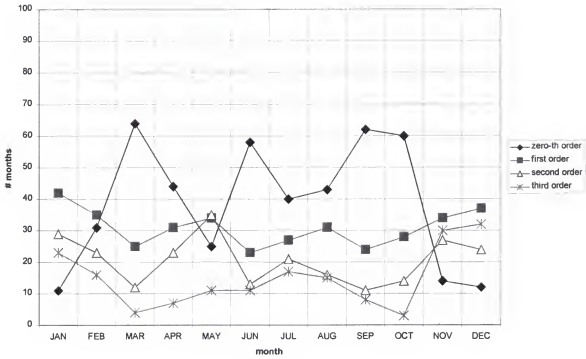


Figure 3-2. Monthly AIC and BIC statistics, El Niño years

AIC -- non-El Niño Years



BIC -- non-El Niño Years

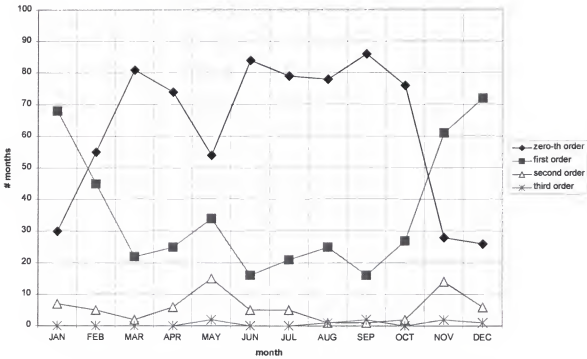


Figure 5-3. Monthly AIC and BIC statistics, non-El Niño years

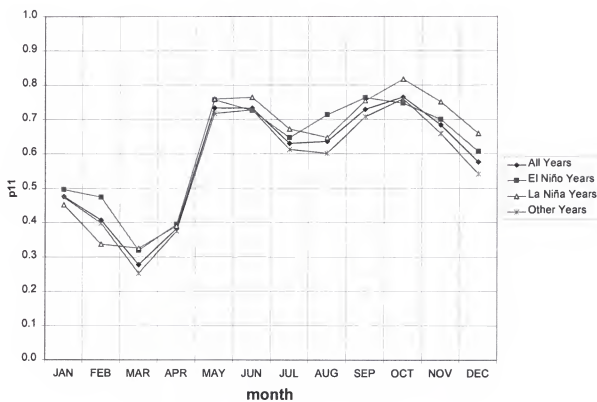
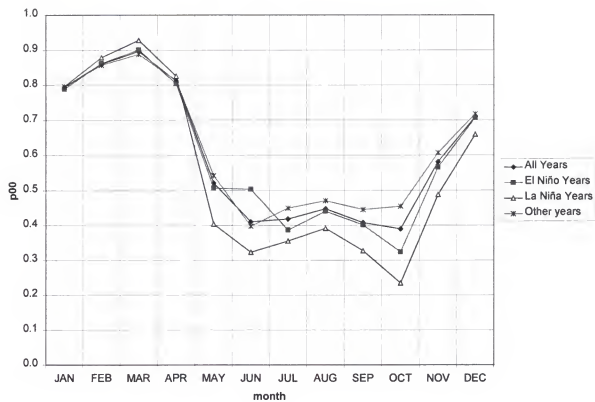


Figure 5-4. Sanatorio Duran p_{00} (top) and p_{11} (bottom)

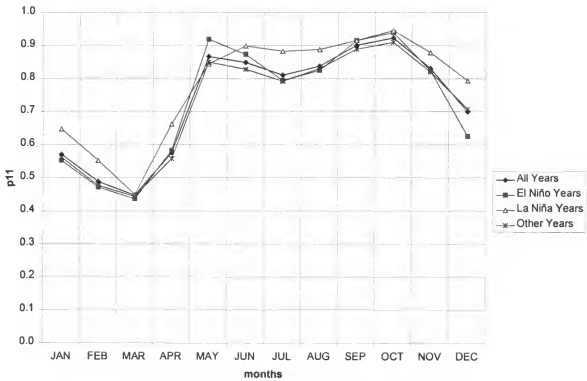
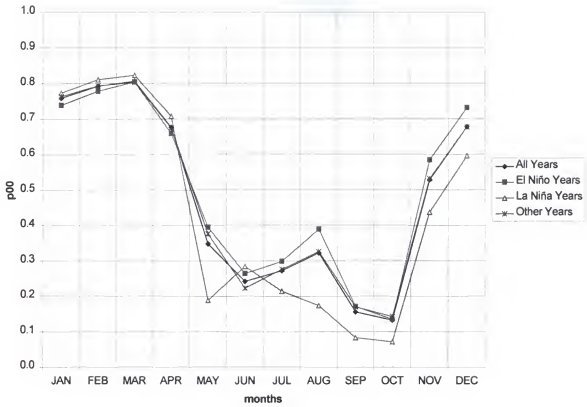
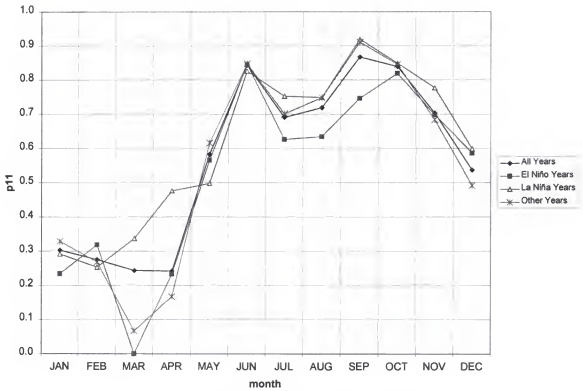
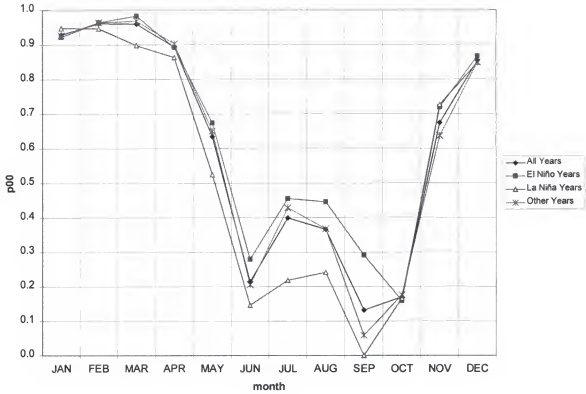
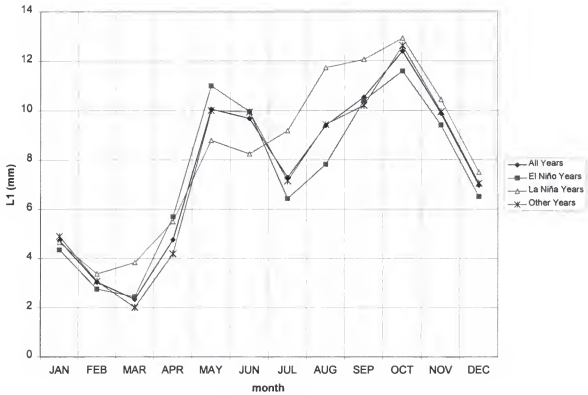


Figure 5-5. Villa Mills p_{00} (top) and p_{11} (bottom)

Figure 5-6. Nagatac p_{00} (top) and p_{11} (bottom)

Figure 5-7. Sanatorio Duran L₁

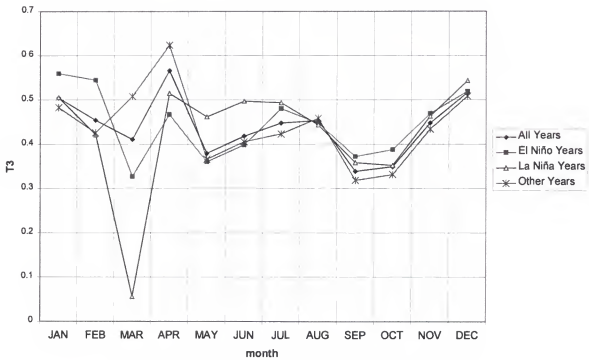
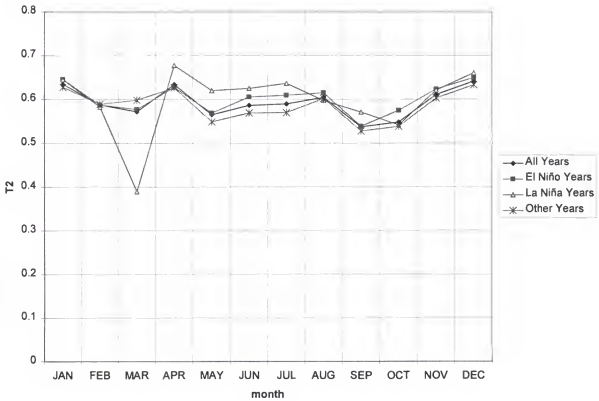
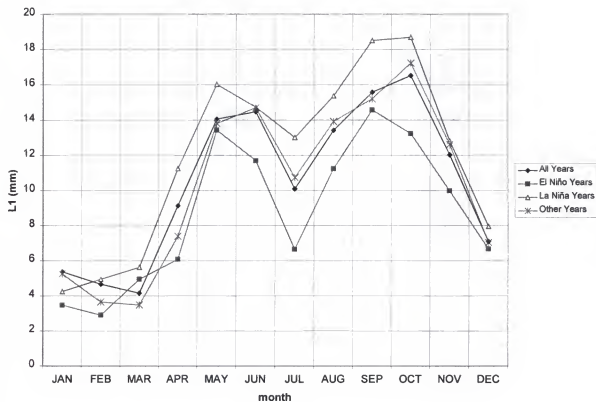
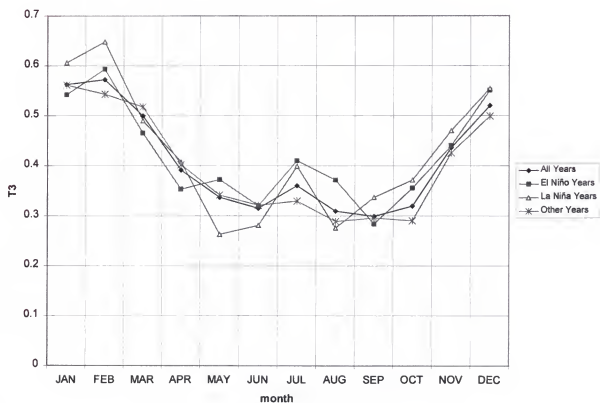
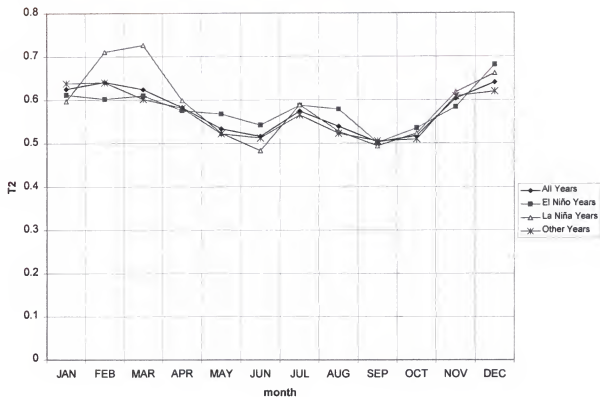
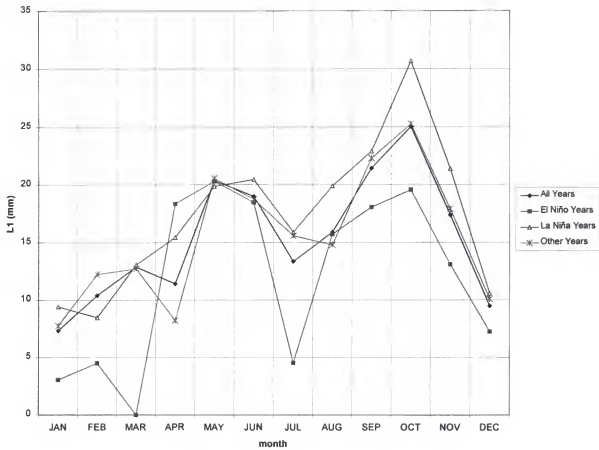
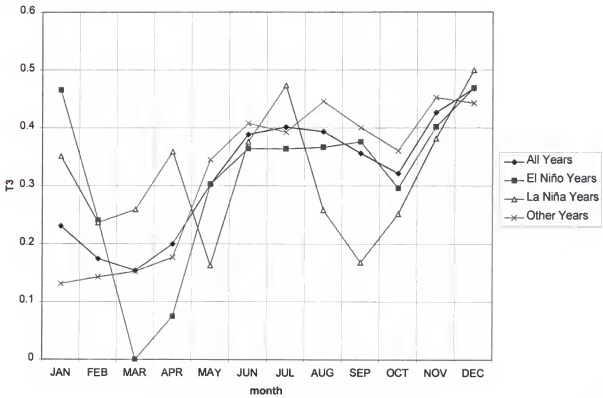
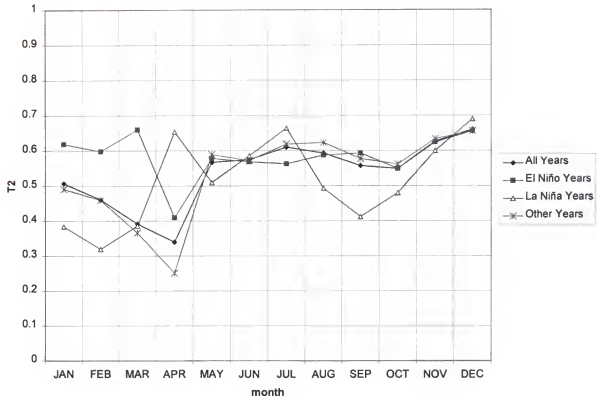


Figure 5-8. Sanatorio Duran T_2 (top) and T_3 (bottom)

Figure 5-9. Villa Mills L₁

Figure 5-10. Villa Mills T_2 (top) and T_3 (bottom)

Figure 5-11. Nagatac L_1

Figure 5-12. Nagatac T_2 (top) and T_3 (bottom)

Sanatorio Duran (073011) Lag Correlations (p00 vs. SOI)

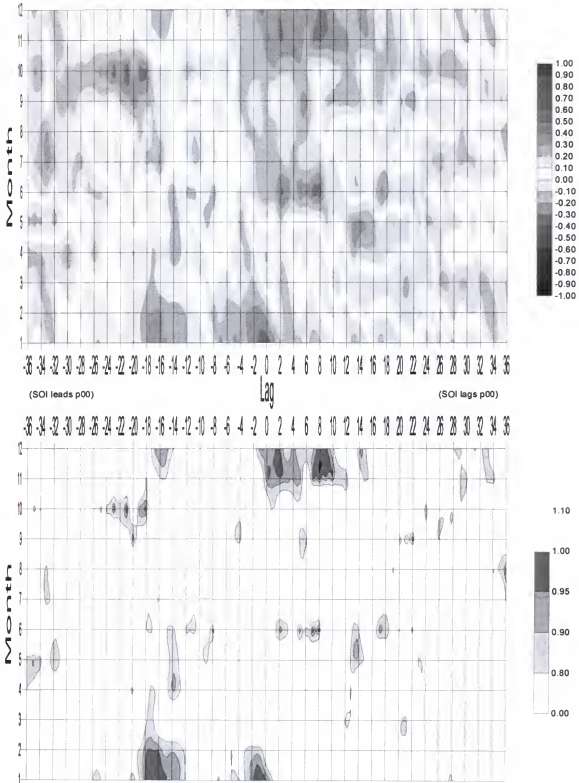


Figure 5-13. Sanatorio Duran lag-correlations: p00 vs. SOI

Sanatorio Duran (073011) Lag Correlations (p11 vs. SOI)

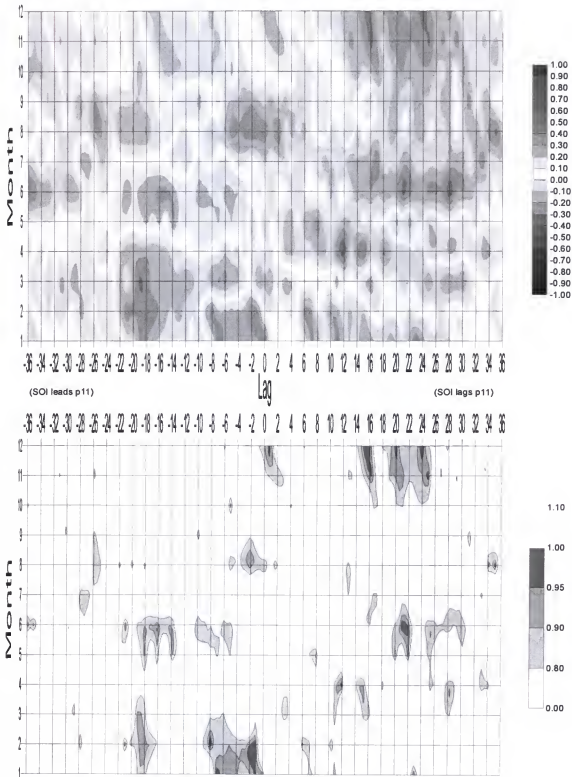


Figure 5-14. Sanatorio Duran lag-correlations: p₁₁ vs. SOI

Sanatorio Duran (073011) Lag Correlations
(p00 vs. NAO)

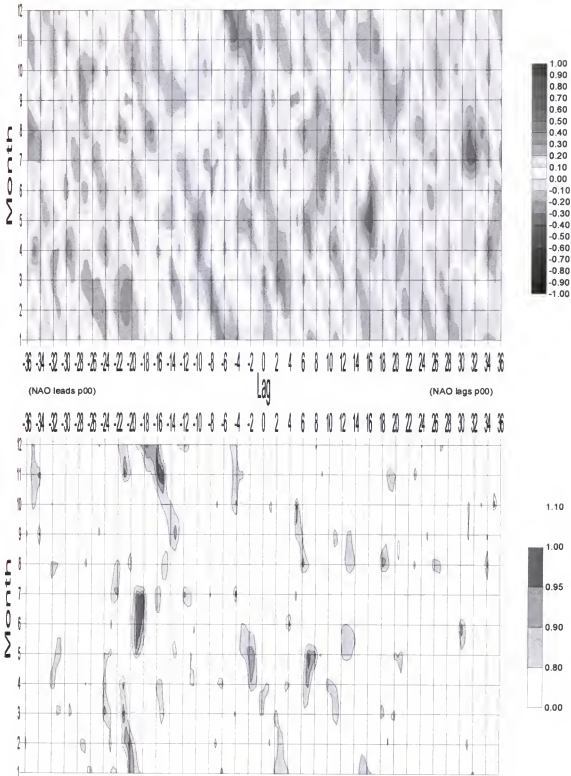


Figure 5-15. Sanatorio Duran lag-correlations: p_{00} vs. NAO

Sanatorio Duran (073011) Lag Correlations (p11 vs. NAO)

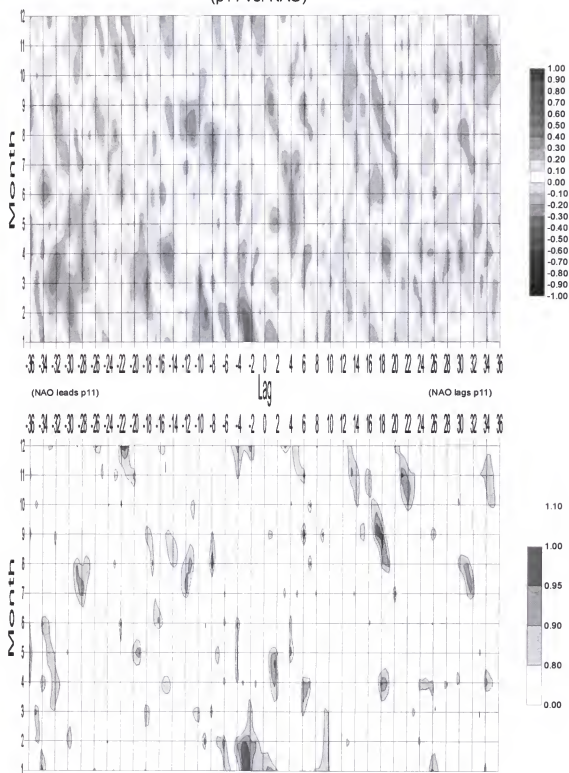


Figure 5-16. Sanatorio Duran lag-correlations: p_{11} vs. NAO

Sanatorio Duran (073011) Lag Correlations (p00 vs. NINO34)

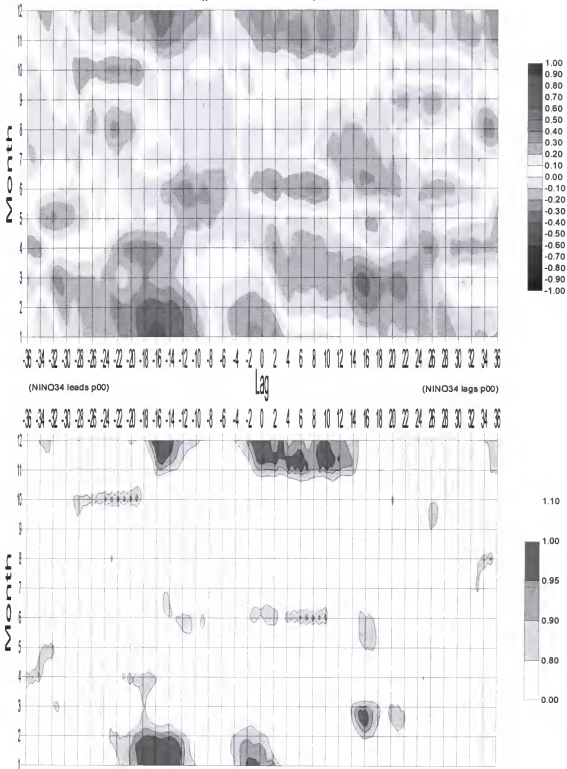


Figure 5-17. Sanatorio Duran lag-correlations: p00 vs. NINO34

Sanatorio Duran (073011) Lag Correlations (p11 vs. NINO34)

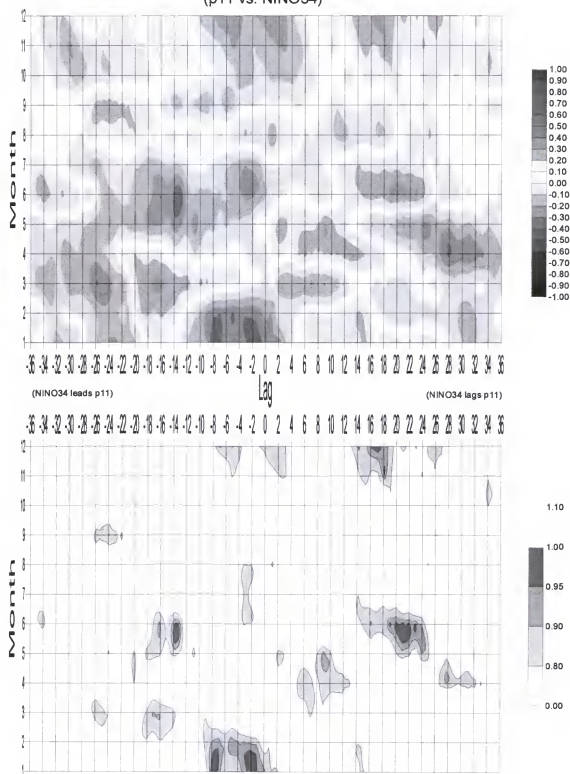


Figure 5-18. Sanatorio Duran lag-correlations: p₁₁ vs. NINO34

Sanatorio Duran (073011) Lag Correlations (p00 vs. QBO50)

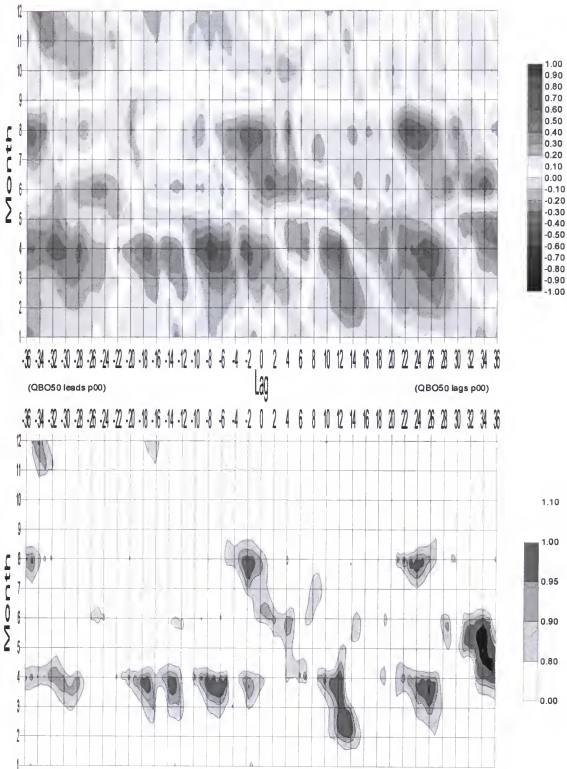


Figure 5-19. Sanatorio Duran lag-correlations: p₀₀ vs. QBO50

Sanatorio Duran (073011) Lag Correlations
(p11 vs. QBO50)

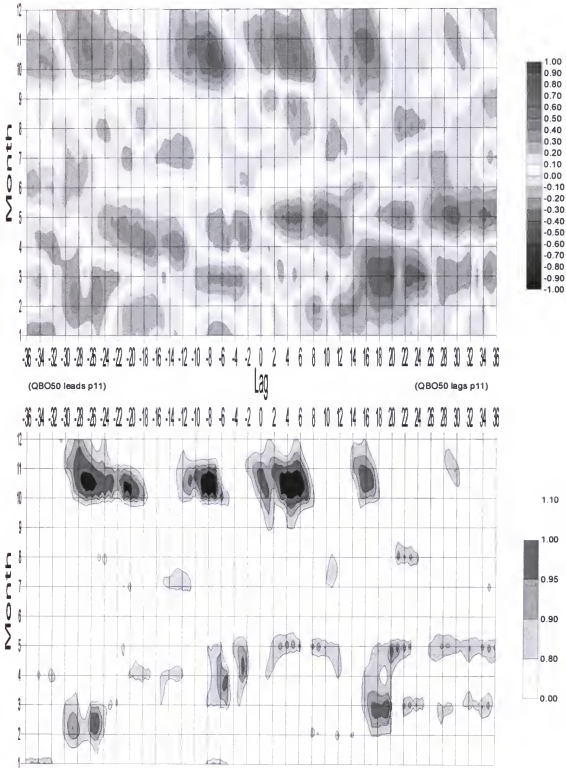


Figure 5-20. Sanatorio Duran lag-correlations: p11 vs. QBO50

Villa Mills (073033) Lag Correlations
(p00 vs. SOI)

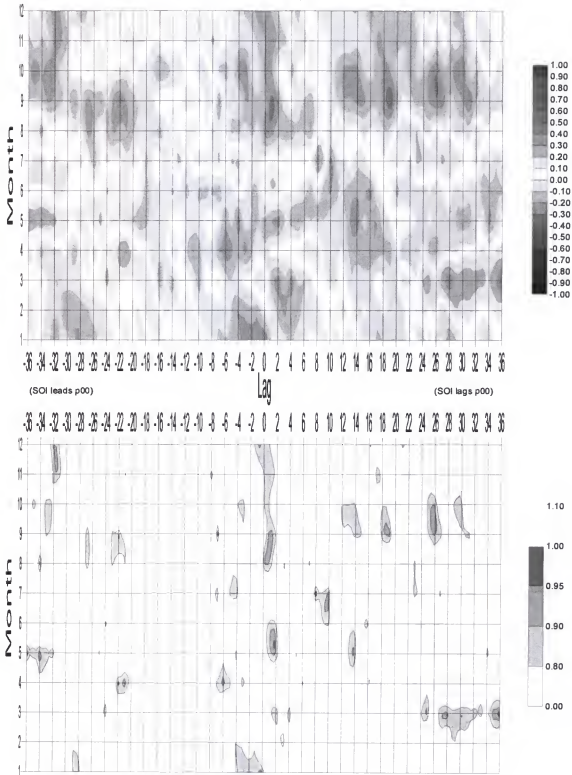


Figure 5-21. Villa Mills lag-correlations: p₀₀ vs. SOI

Villa Mills (073033) Lag Correlations
(p11 vs. SOI)

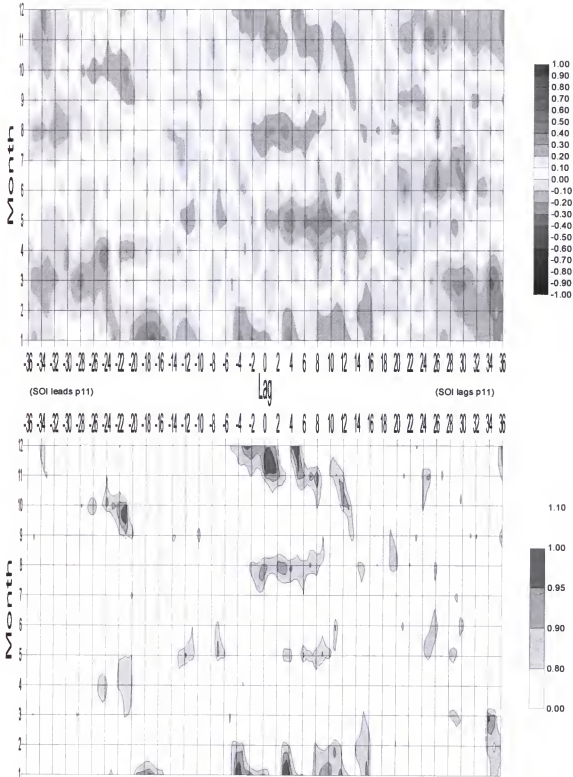


Figure 5-22. Villa Mills lag-correlations: p11 vs. SOI

Villa Mills (073033) Lag Correlations (p00 vs. NAO)

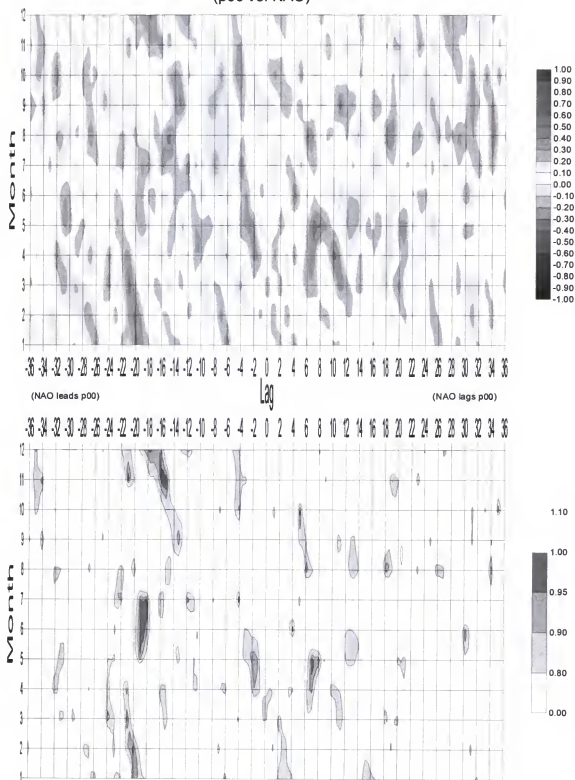


Figure 5-23. Villa Mills lag-correlations: p00 vs. NAO

Villa Mills (073033) Lag Correlations
(p11 vs. NAO)

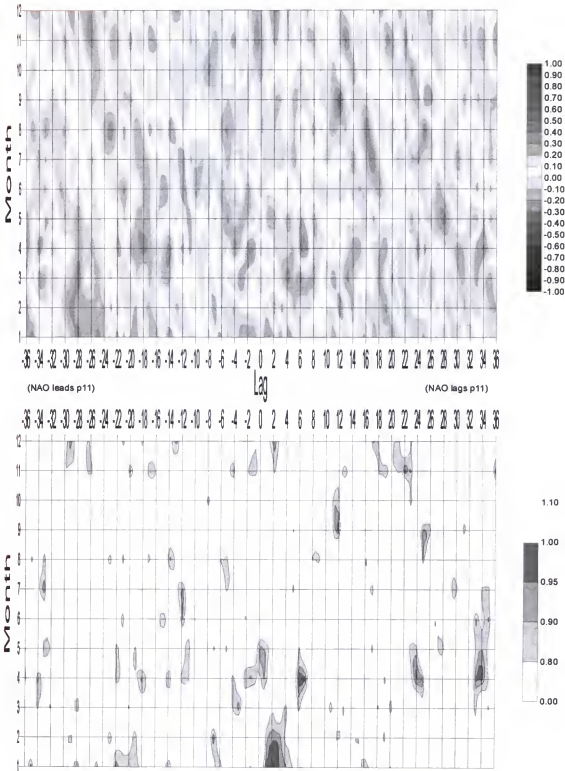


Figure 5-24. Villa Mills lag-correlations: p₁₁ vs. NAO

Villa Mills (073033) Lag Correlations
(p00 vs. NINO34)

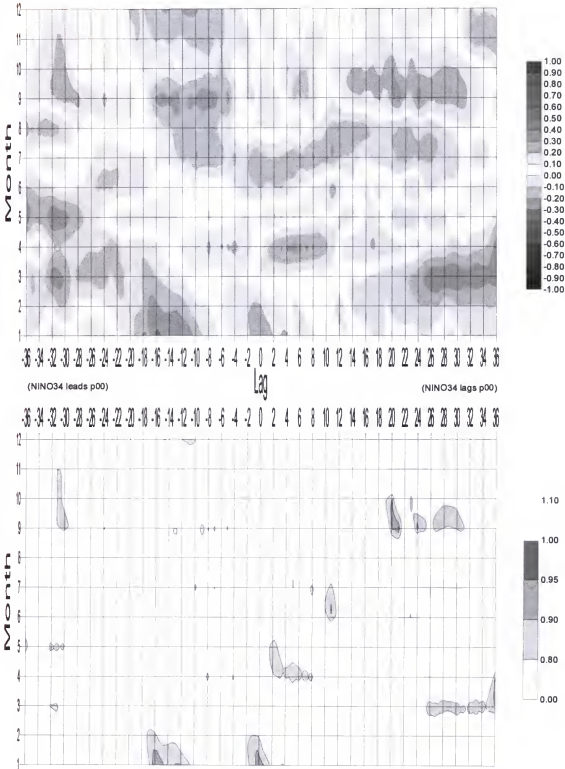


Figure 5-25. Villa Mills lag-correlations: p₀₀ vs. NINO34

Villa Mills (073033) Lag Correlations
(p11 vs. NINO34)

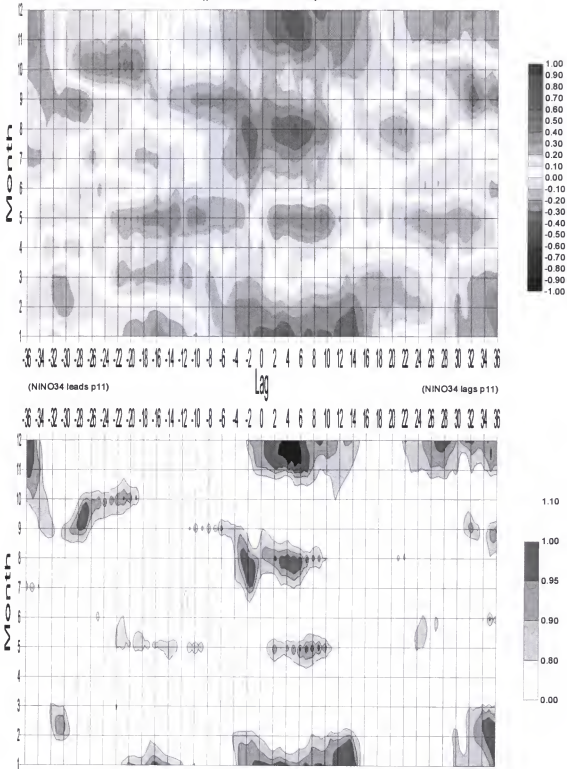


Figure 5-26. Villa Mills lag-correlations: p₁₁ vs. NINO34

Villa Mills (073033) Lag Correlations
(p00 vs. QBO50)

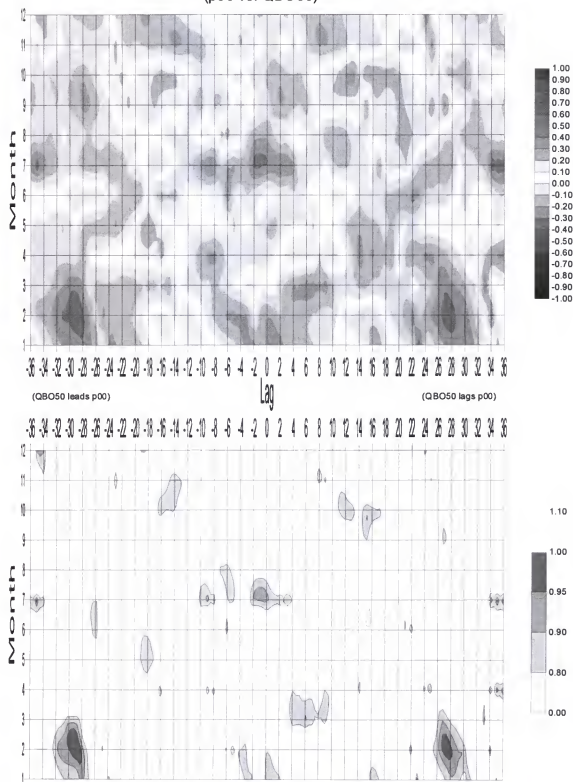


Figure 5-27. Villa Mills lag-correlations: p_{00} vs. QBO50

Villa Mills (073033) Lag Correlations (p11 vs. QBO50)

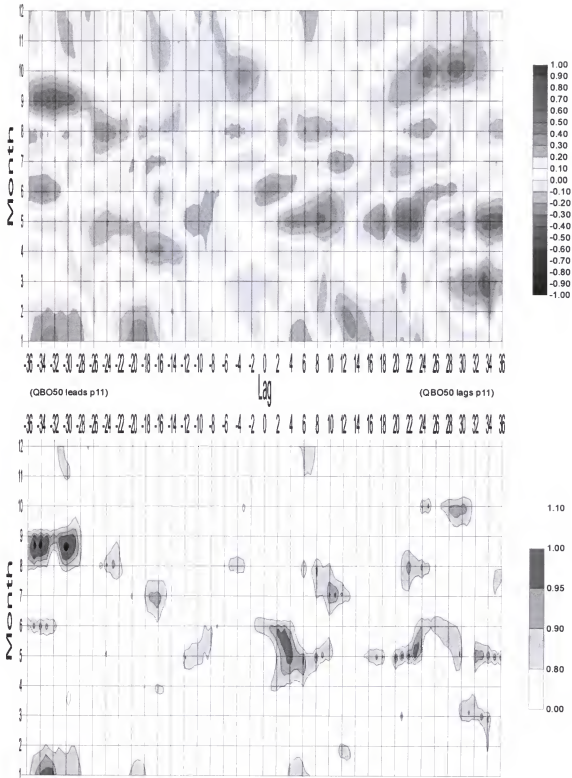


Figure 5-28. Villa Mills lag-correlations: p₁₁ vs. QBO50

Nagatac (080005) Lag Correlations (p00 vs. SOI)

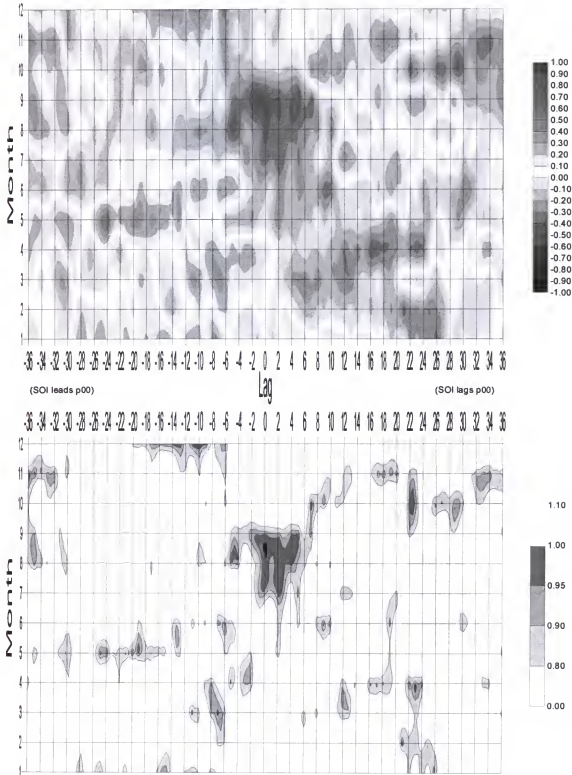


Figure 5-29. Nagatac lag-correlations: p₀₀ vs. SOI

Nagatac (080005) Lag Correlations (p11 vs. SOI)

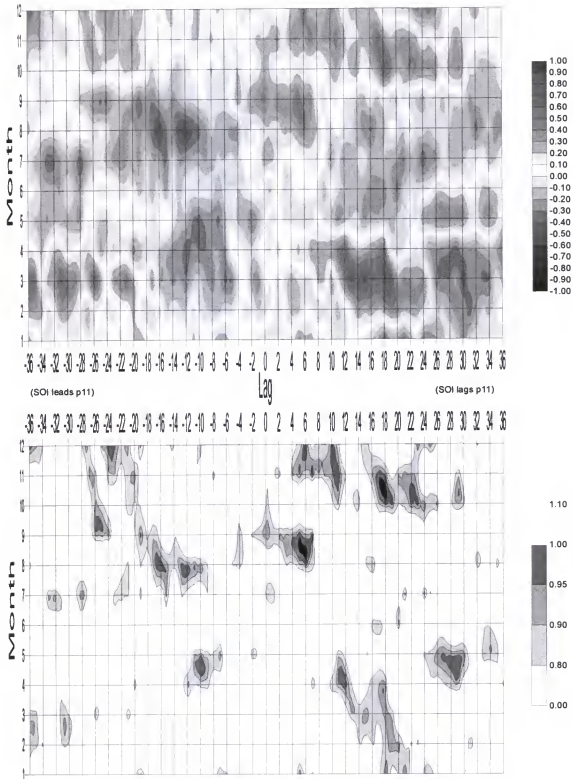


Figure 5-30. Nagatac lag-correlations: p_{11} vs. SOI

Nagatac (080005) Lag Correlations (p00 vs. NAO)

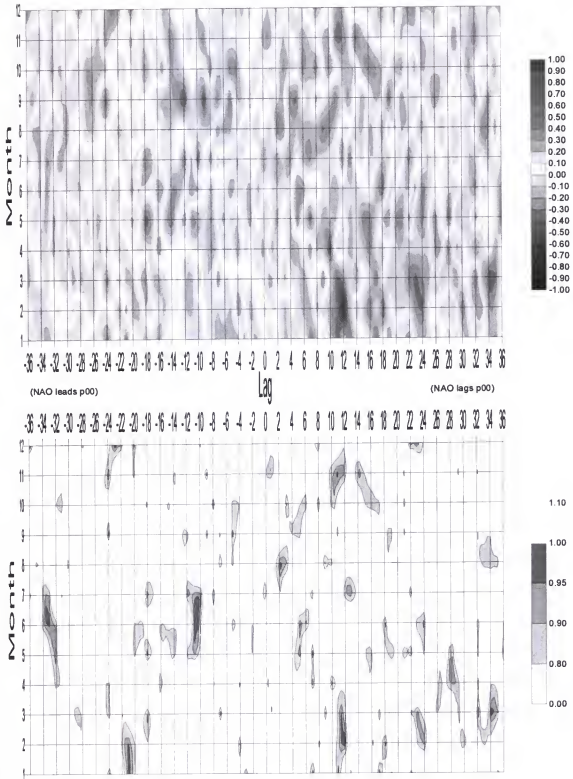


Figure 5-31. Nagatac lag-correlations: p₀₀ vs. NAO

Nagatac (080005) Lag Correlations (p11 vs. NAO)

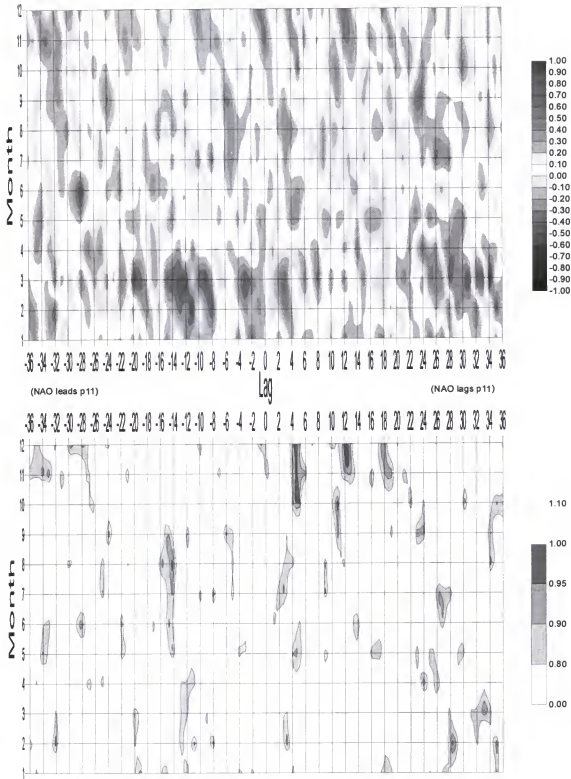


Figure 5-32. Nagatac lag-correlations: p₁₁ vs. NAO

Nagatac (080005) Lag Correlations (p00 vs. NINO34)

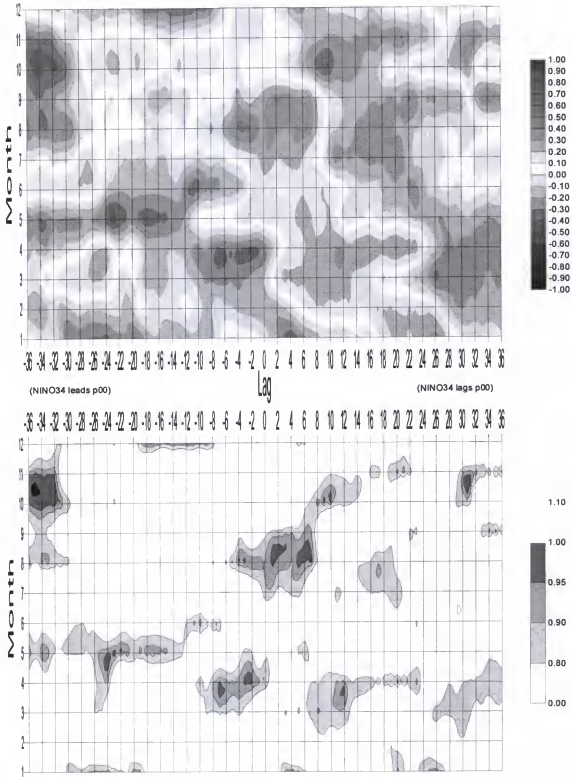


Figure 5-33. Nagatac lag-correlations: p00 vs. NINO34

Nagatac (080005) Lag Correlations (p11 vs. NINO34)

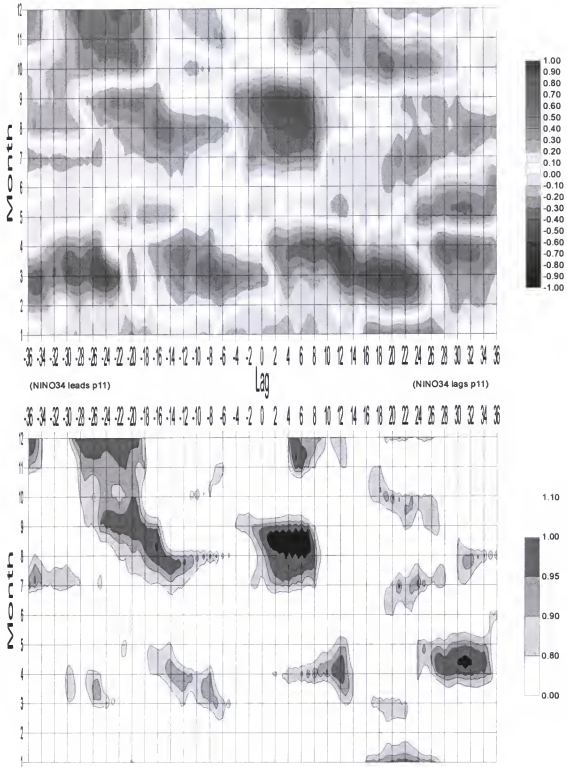


Figure 5-34. Nagatac lag-correlations: p₁₁ vs. NINO34

Nagatac (080005) Lag Correlations (p00 vs. QBO50)

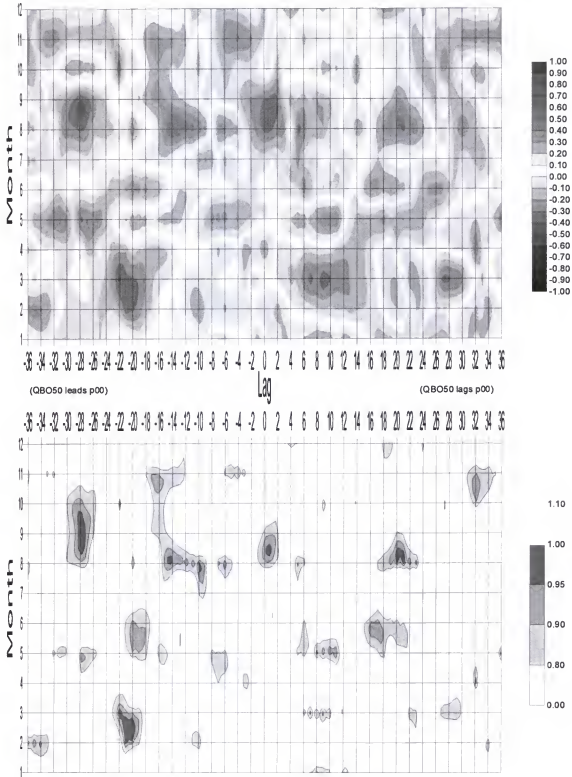


Figure 5-35. Nagatac lag-correlations: p00 vs. QBO50

Nagatac (080005) Lag Correlations (p11 vs. QBO50)

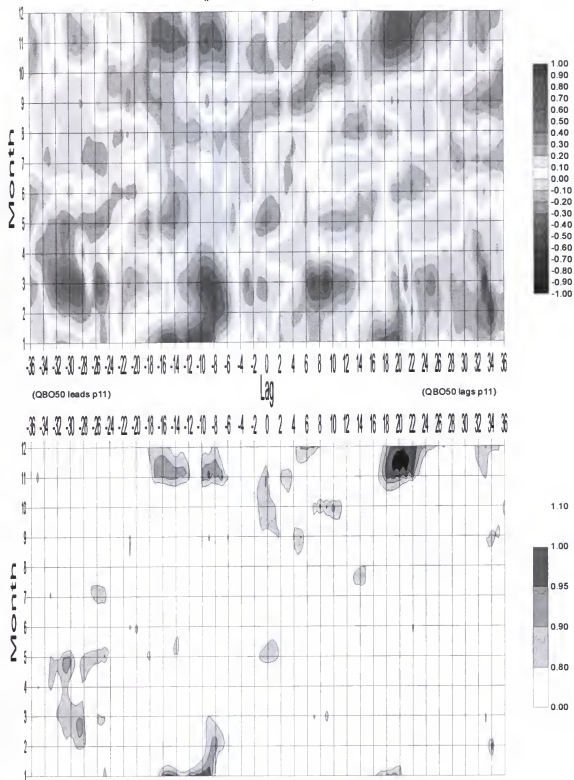


Figure 5-36. Nagatac lag-correlations: p11 vs. QBO50

Sanatorio Duran (073011) Lag Correlations (L1 vs. SOI)

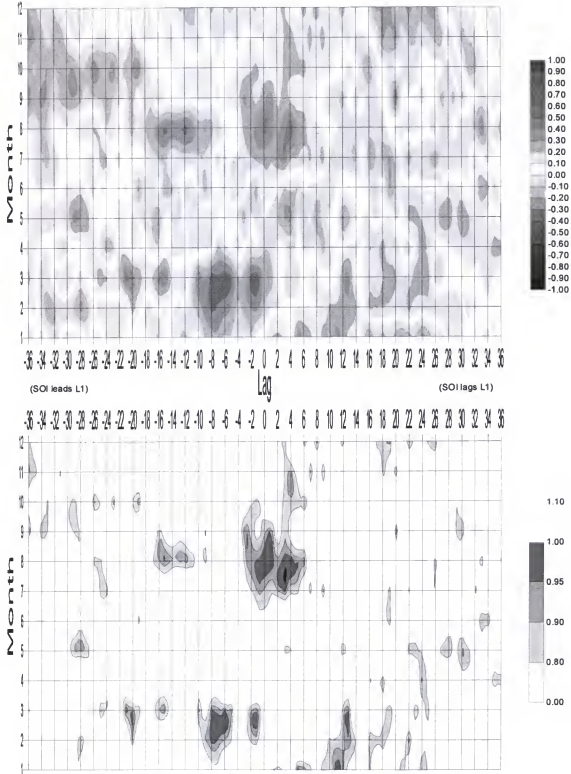


Figure 5-37. Sanatorio Duran lag-correlations: L₁ vs. SOI

Sanatorio Duran (073011) Lag Correlations (T₂ vs. SOI)

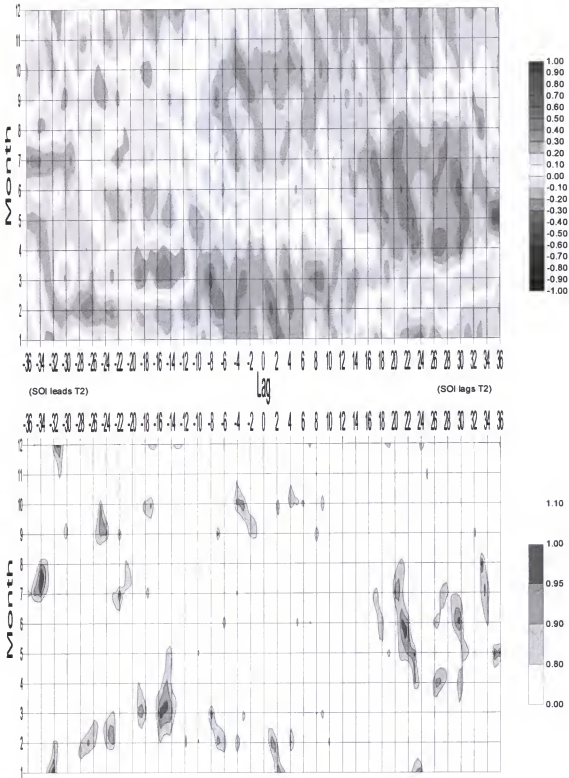


Figure 5-38. Sanatorio Duran lag-correlations: T₂ vs. SOI

Sanatorio Duran (073011) Lag Correlations (T3 vs. SOI)

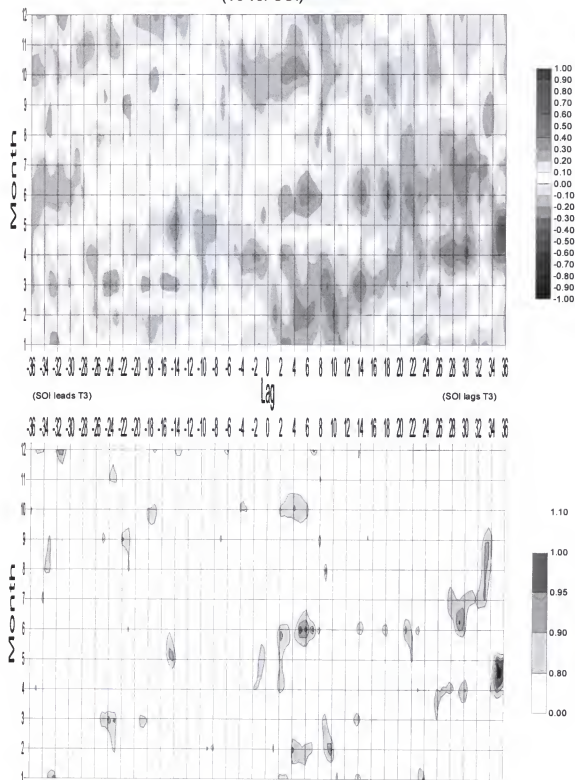


Figure 5-39. Sanatorio Duran lag-correlations: T₃ vs. SOI

Sanatorio Duran (073011) Lag Correlations (L1 vs. NAO)

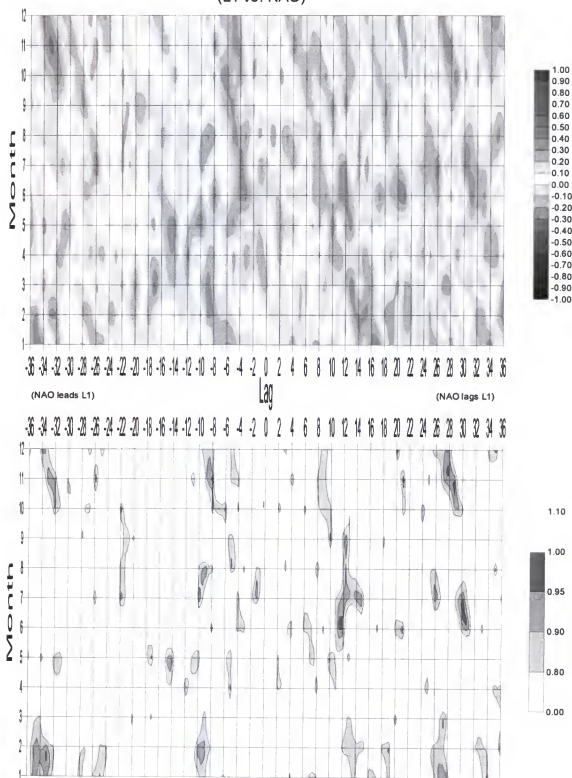


Figure 5-40. Sanatorio Duran lag-correlations: L₁ vs. NAO

Sanatorio Duran (073011) Lag Correlations (T₂ vs. NAO)

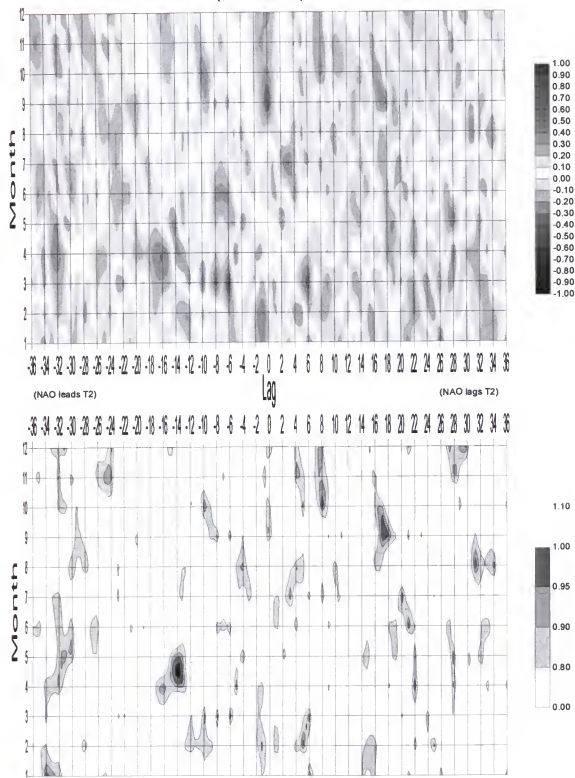


Figure 5-41. Sanatorio Duran lag-correlations: T₂ vs. NAO

Sanatorio Duran (073011) Lag Correlations (T₃ vs. NAO)

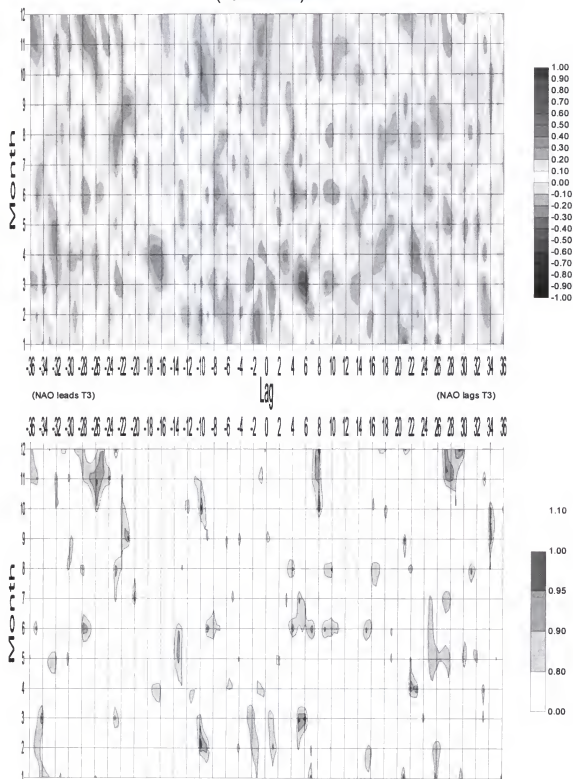


Figure 5-42. Sanatorio Duran lag-correlations: T₃ vs. NAO

Sanatorio Duran (073011) Lag Correlations (L1 vs. NINO34)

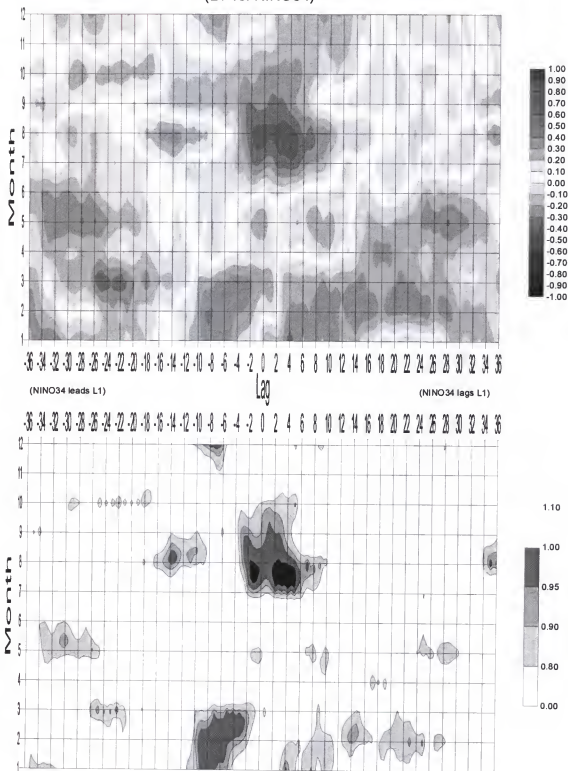


Figure 5-43. Sanatorio Duran lag-correlations: L₁ vs. NINO34

Sanatorio Duran (073011) Lag Correlations
(T₂ vs. NINO34)

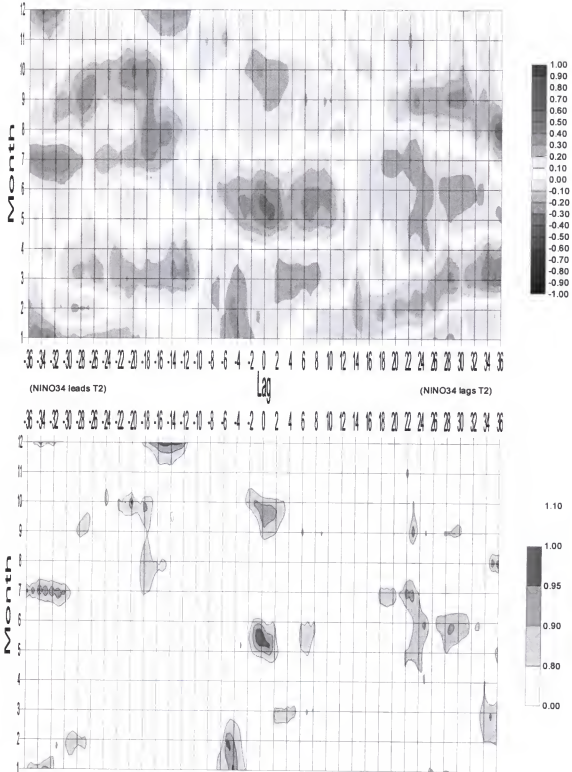


Figure 5-44. Sanatorio Duran lag-correlations: T₂ vs. NINO34

Sanatorio Duran (073011) Lag Correlations (T₃ vs. NINO34)

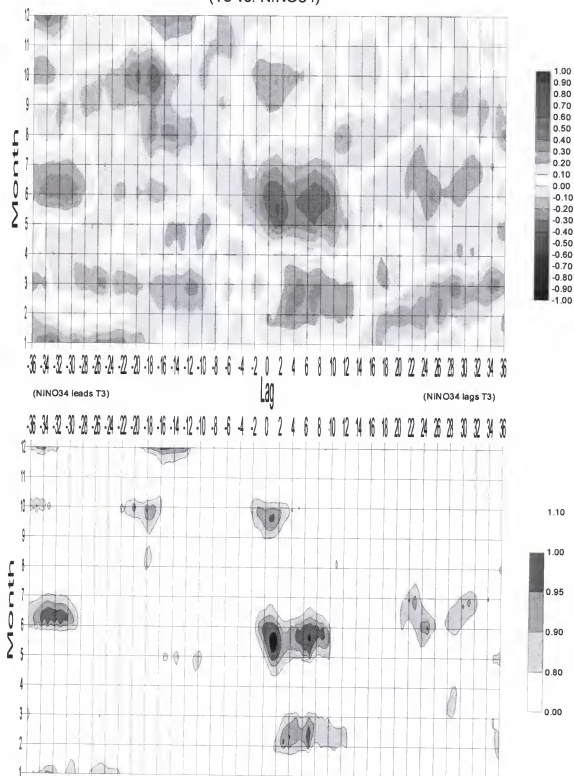


Figure 5-45. Sanatorio Duran lag-correlations: T₃ vs. NINO34

Sanatorio Duran (073011) Lag Correlations (L1 vs. QBO50)

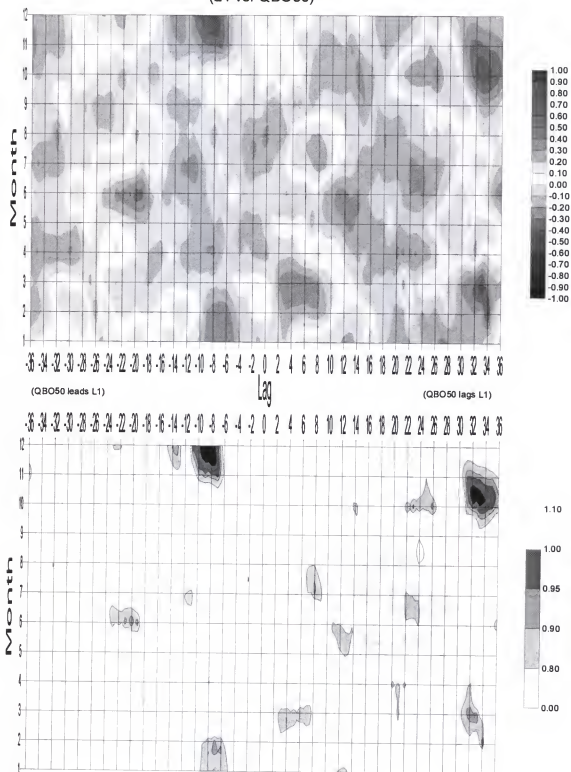


Figure 5-46. Sanatorio Duran lag-correlations: L_1 vs. QBO50

Sanatorio Duran (073011) Lag Correlations (T₂ vs. QBO50)

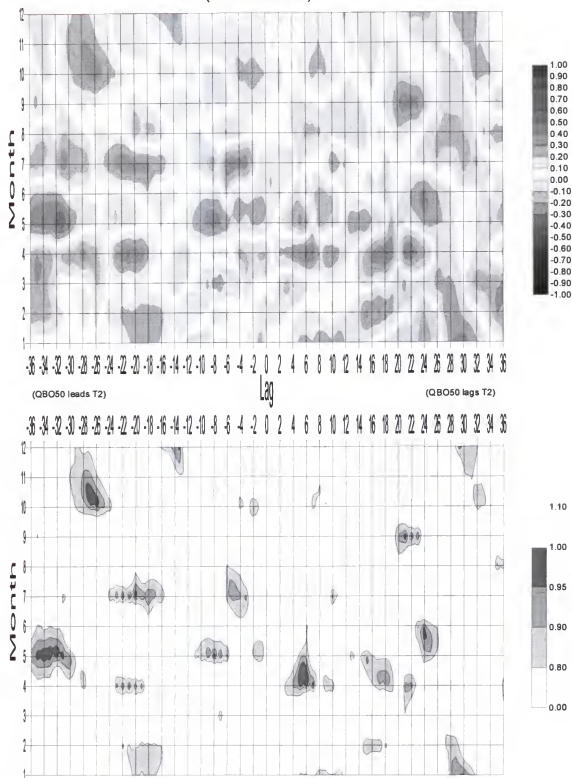


Figure 5-47. Sanatorio Duran lag-correlations: T₂ vs. QBO50

Sanatorio Duran (073011) Lag Correlations (T3 vs. QBO50)

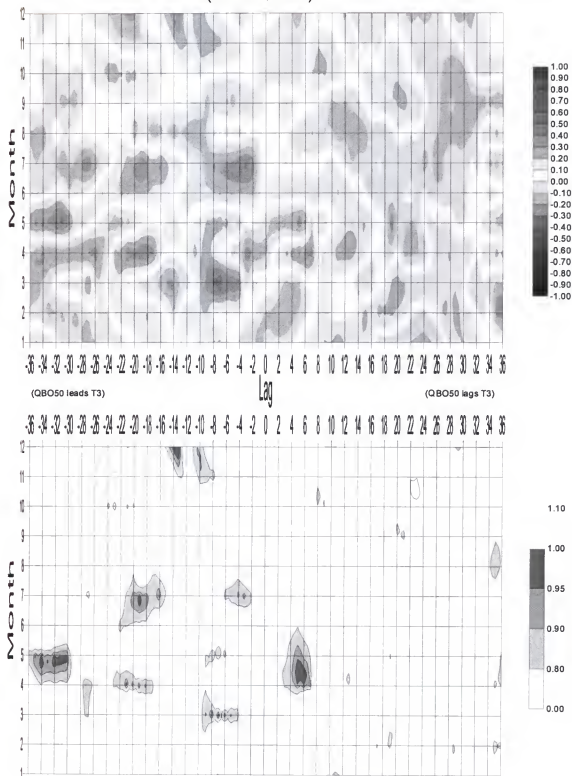


Figure 5-48. Sanatorio Duran lag-correlations: T₃ vs. QBO50

Villa Mills (073033) Lag Correlations (L1 vs. SOI)

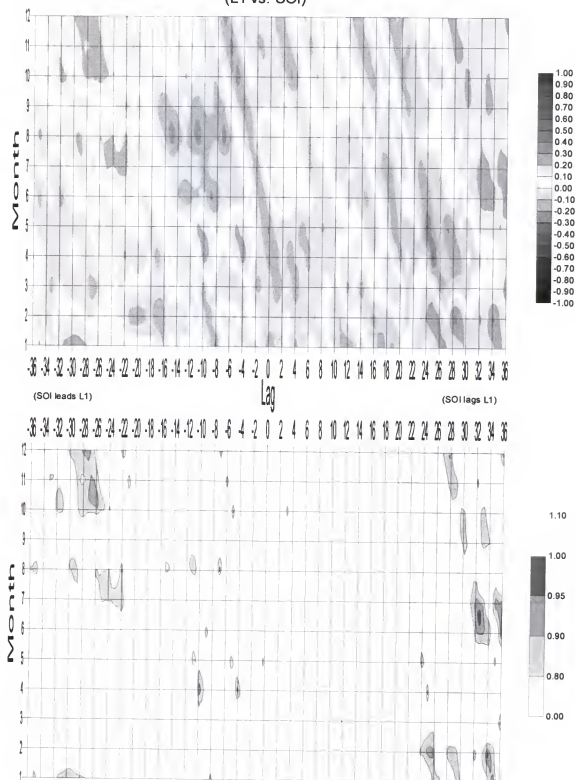


Figure 5-49. Villa Mills lag-correlation: L₁ vs. SOI

Villa Mills (073033) Lag Correlations (T2 vs. SOI)

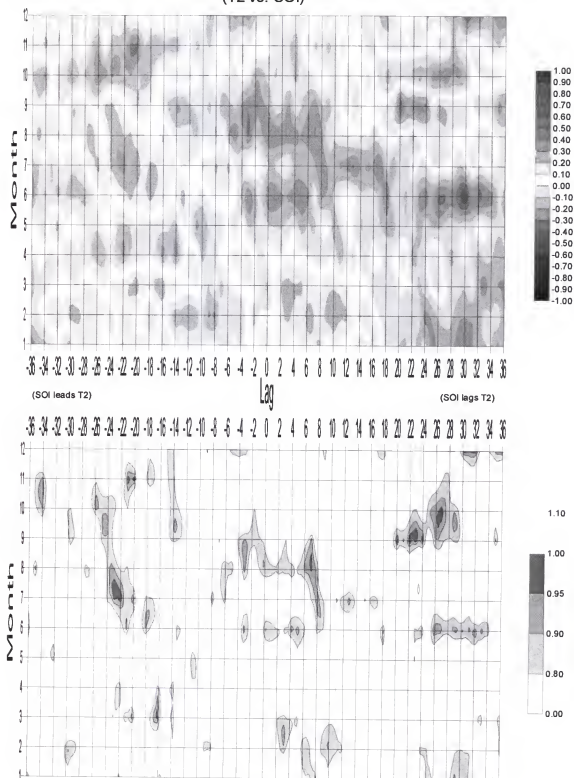


Figure 5-50. Villa Mills lag-correlation: T2 vs. SOI

Villa Mills (073033) Lag Correlations (T3 vs. SOI)

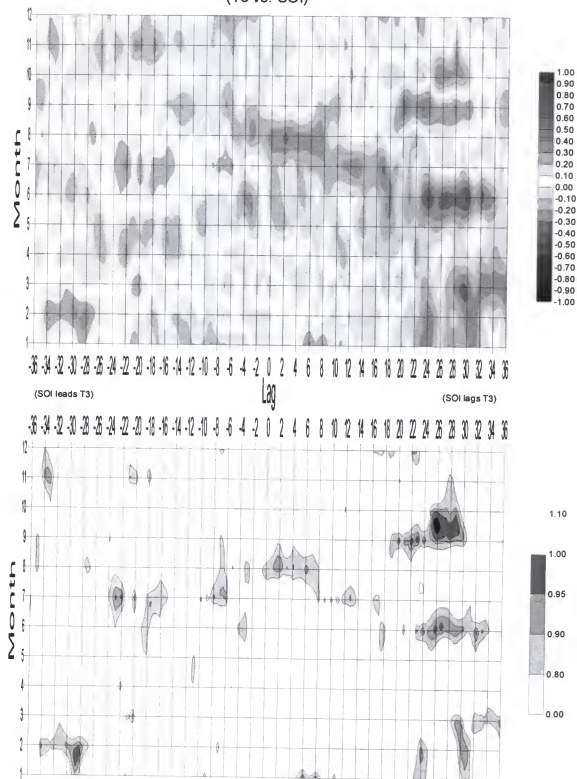


Figure 5-51. Villa Mills lag-correlation: T₃ vs. SOI

Villa Mills (073033) Lag Correlations (L1 vs. NAO)

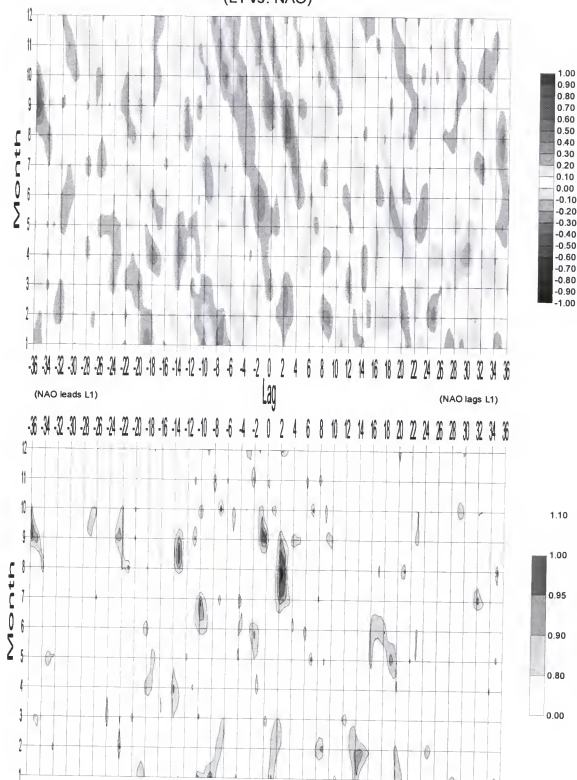


Figure 5-52. Villa Mills lag-correlation: L₁ vs. NAO

Villa Mills (073033) Lag Correlations (T₂ vs. NAO)

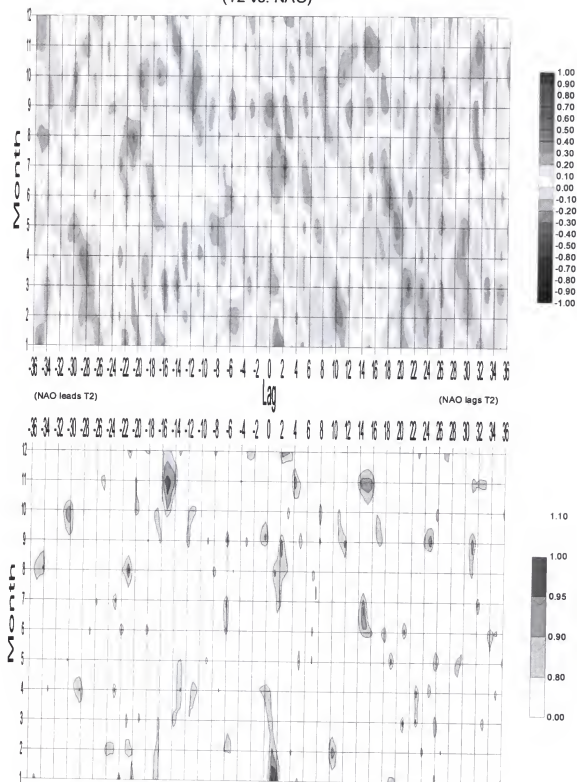


Figure 5-53. Villa Mills lag-correlation: T₂ vs. NAO

Villa Mills (073033) Lag Correlations (T3 vs. NAO)

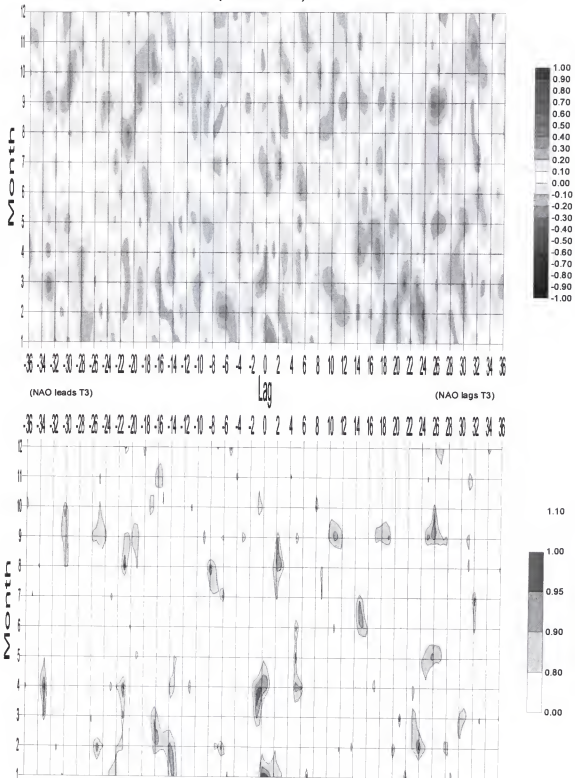


Figure 5-54. Villa Mills lag-correlation: T₃ vs. NAO

Villa Mills (073033) Lag Correlations (L1 vs. NINO34)

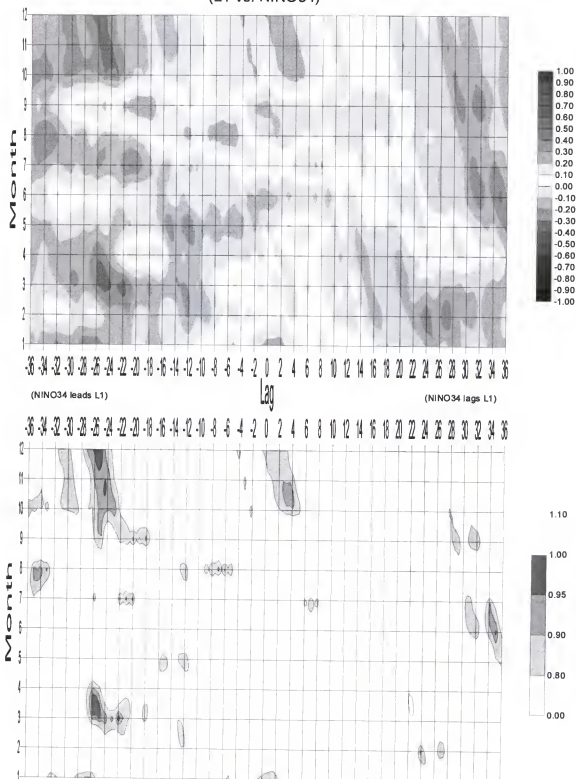


Figure 5-55. Villa Mills lag-correlation: L₁ vs. NINO34

Villa Mills (073033) Lag Correlations
(T₂ vs. NINO34)

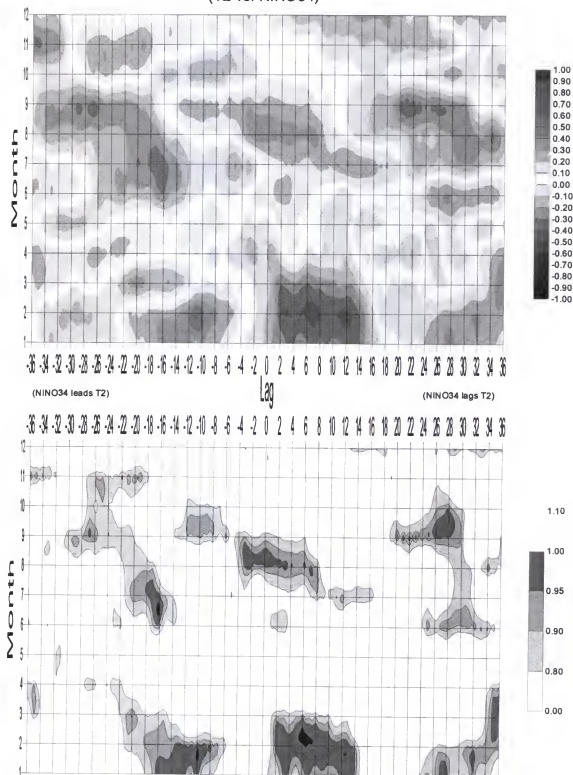


Figure 5-56. Villa Mills lag-correlation: T₂ vs. NINO34

Villa Mills (073033) Lag Correlations
(T₃ vs. NINO34)

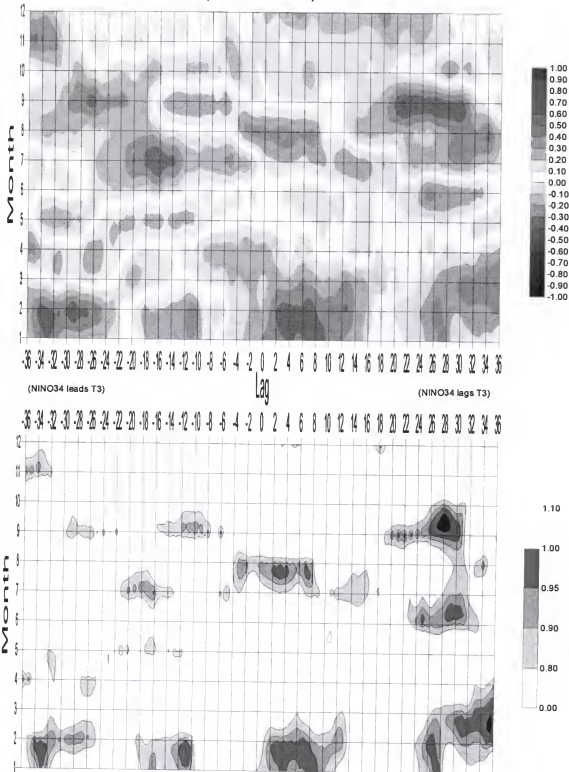


Figure 5-57. Villa Mills lag-correlation: T₃ vs. NINO34

Villa Mills (073033) Lag Correlations (L1 vs. QBO50)

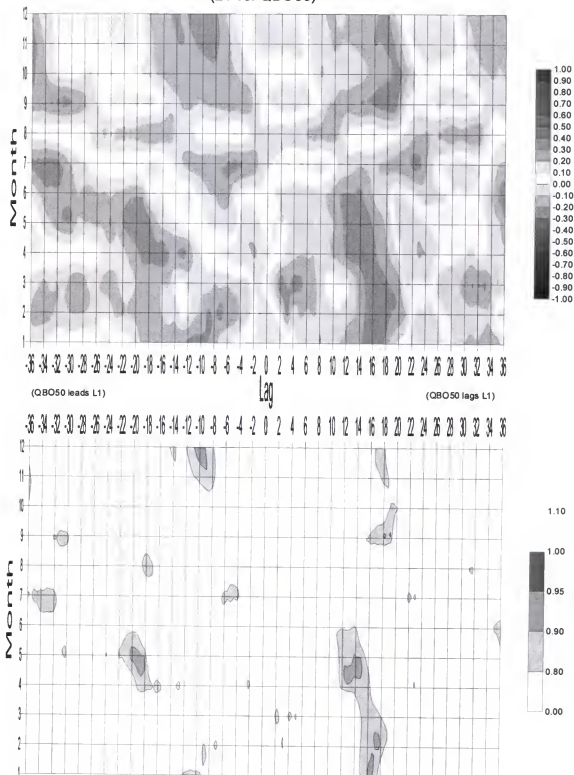


Figure 5-58. Villa Mills lag-correlation: L₁ vs. QBO50

Villa Mills (073033) Lag Correlations
(T₂ vs. QBO50)

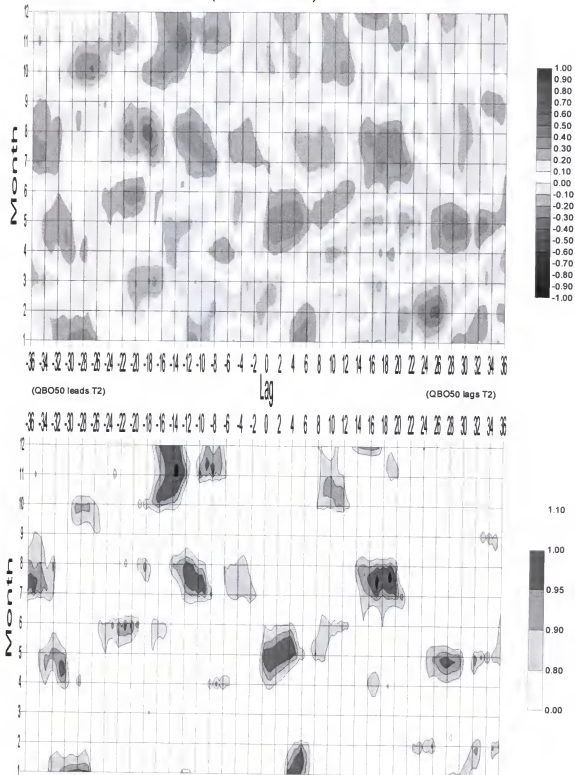


Figure 5-59. Villa Mills lag-correlation: T₂ vs. QBO50

Villa Mills (073033) Lag Correlations
(T₃ vs. QBO50)

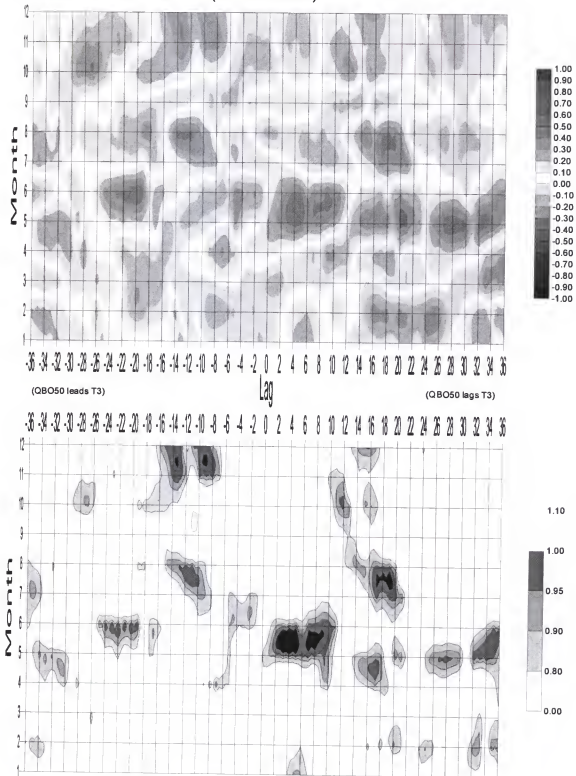


Figure 5-60. Villa Mills lag-correlation: T₃ vs. QBO50

Nagatac (080005) Lag Correlations (L1 vs. SOI)

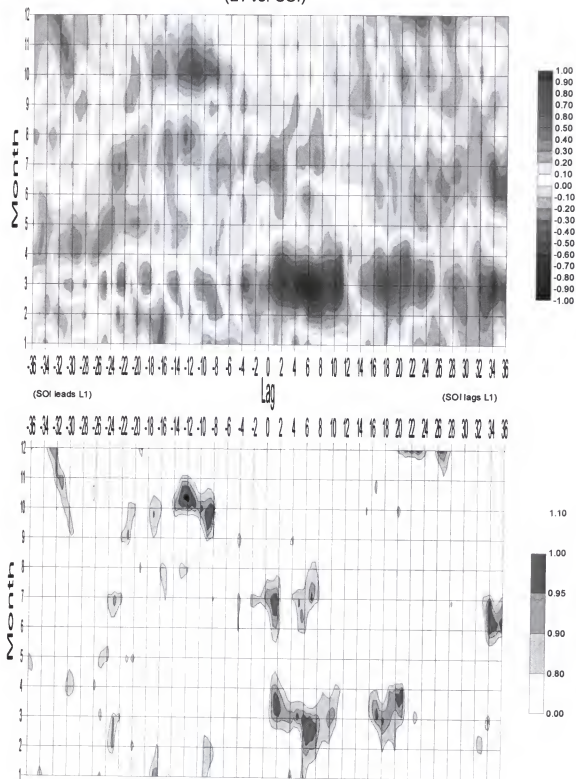


Figure 5-61. Nagatac lag-correlation: L₁ vs. SOI

Nagatac (080005) Lag Correlations (T₂ vs. SOI)

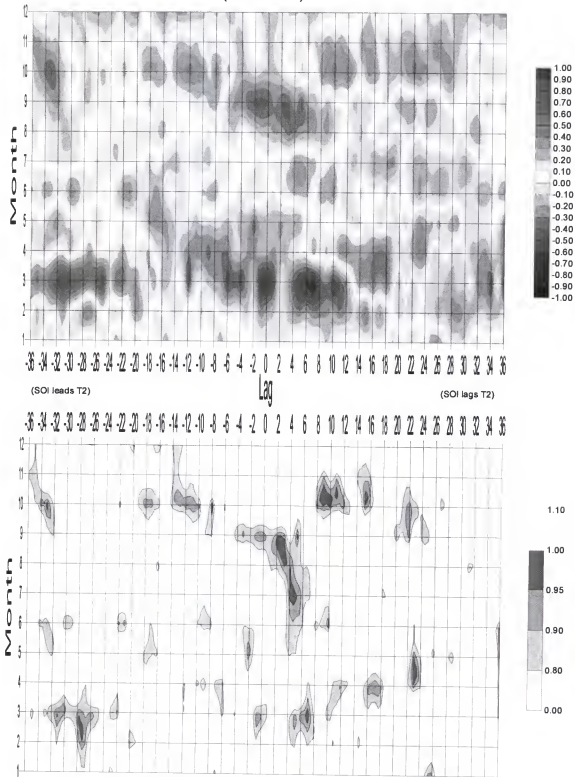


Figure 5-62. Nagatac lag-correlation: T₂ vs. SOI

Nagatac (080005) Lag Correlations (T₃ vs. SOI)

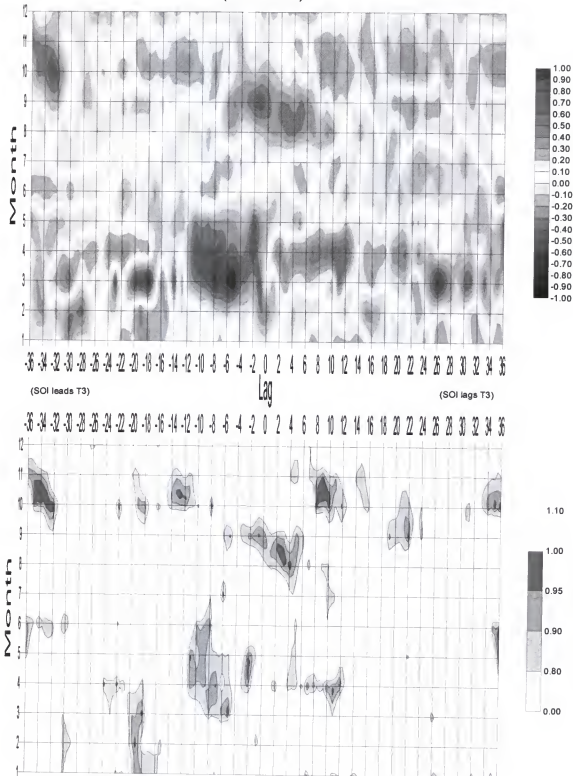


Figure 5-63. Nagatac lag-correlation: T₃ vs. SOI

Nagatac (080005) Lag Correlations (L1 vs. NAO)

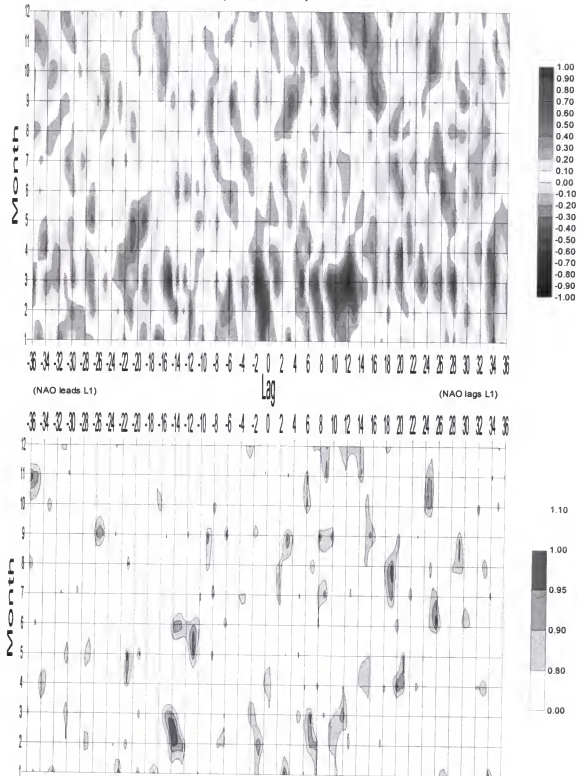


Figure 5-64. Nagatac lag-correlation: L₁ vs. NAO

Nagatac (080005) Lag Correlations (T2 vs. NAO)

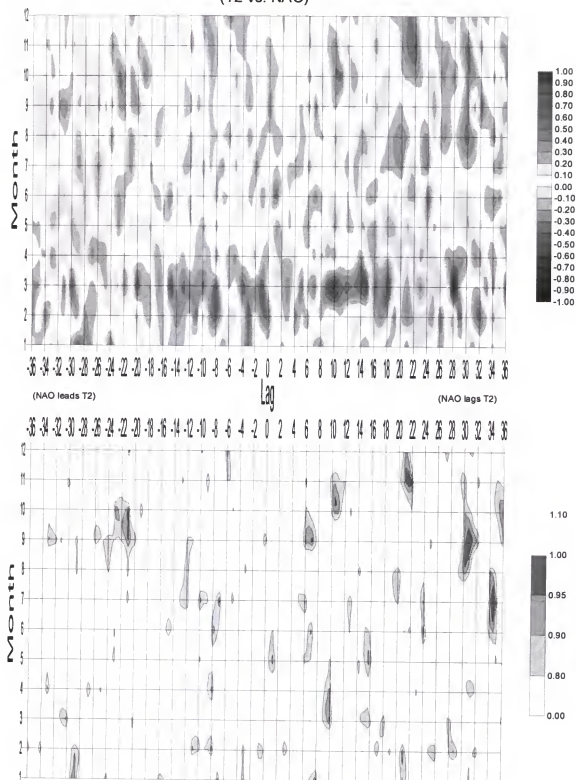


Figure 5-65. Nagatac lag-correlation: T₂ vs. NAO

Nagatac (080005) Lag Correlations (T₃ vs. NAO)

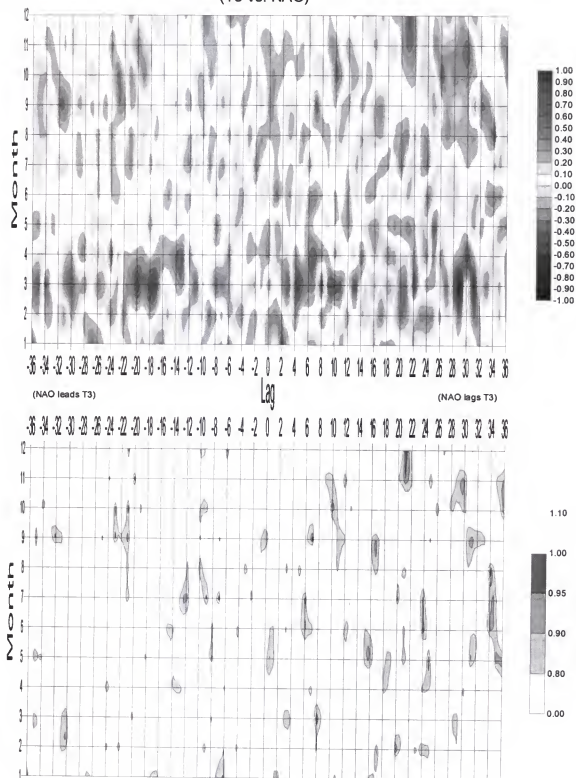


Figure 5-66. Nagatac lag-correlation: T₃ vs. NAO

Nagatac (080005) Lag Correlations (L1 vs. NINO34)

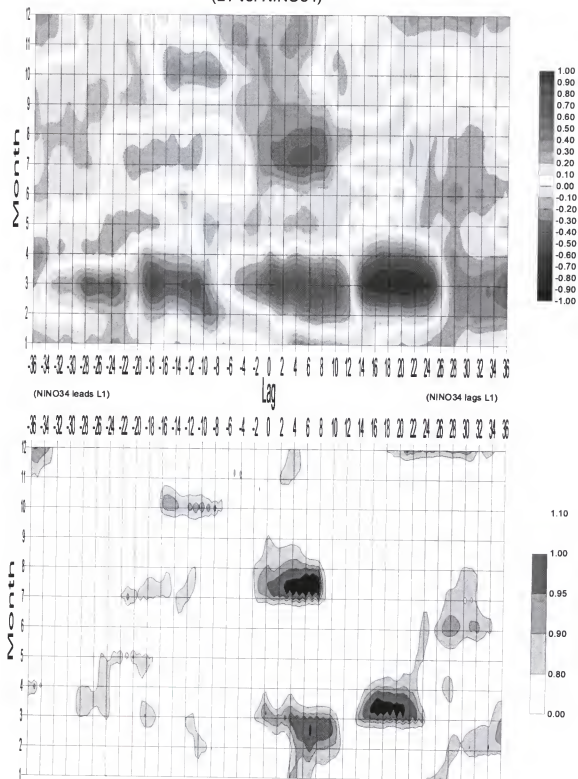


Figure 5-67. Nagatac lag-correlation: L₁ vs. NINO34

Nagatac (080005) Lag Correlations (T₂ vs. NINO34)

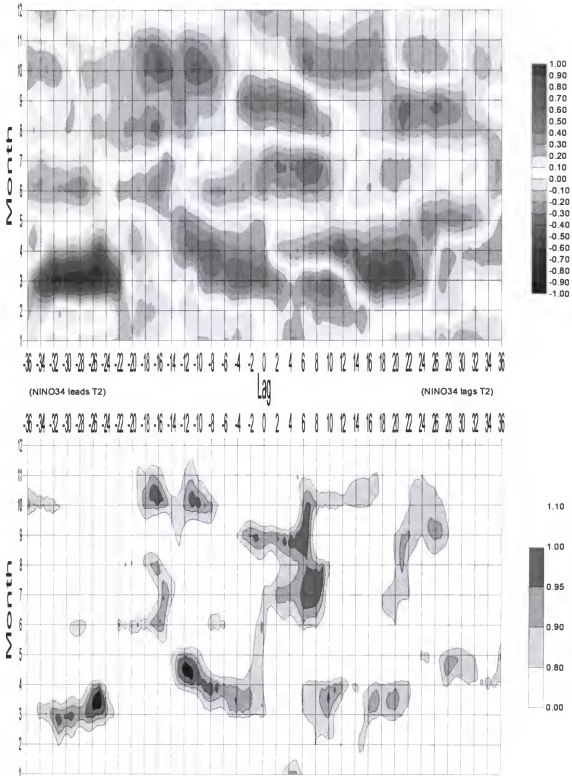


Figure 5-68. Nagatac lag-correlation: T₂ vs. NINO34

Nagatac (080005) Lag Correlations (T3 vs. NINO34)

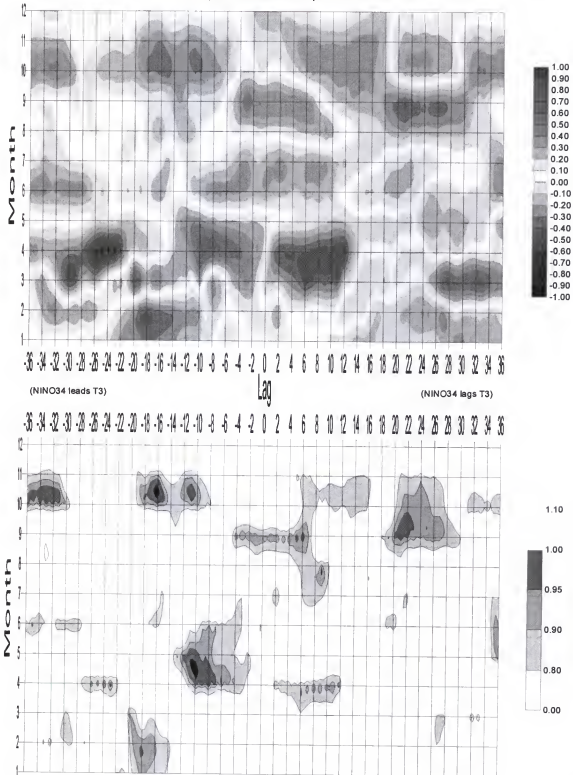


Figure 5-69. Nagatac lag-correlation: T₃ vs. NINO34

Nagatac (080005) Lag Correlations (L1 vs. QBO50)

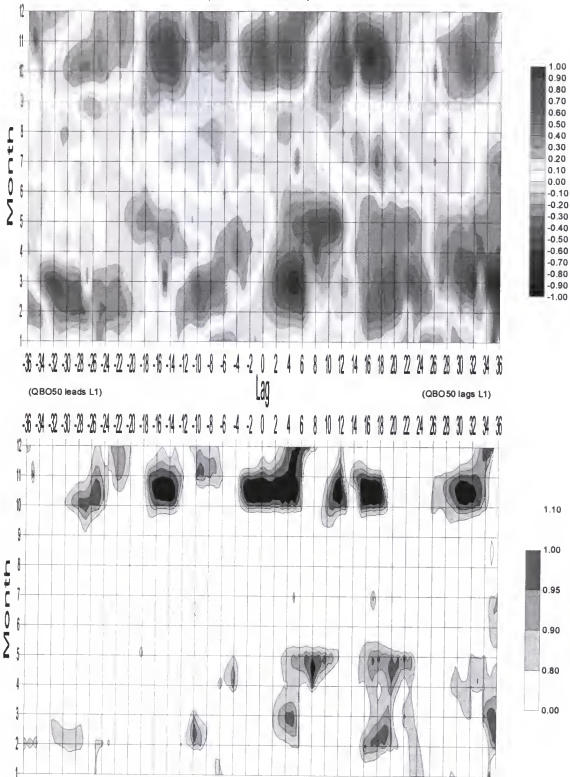


Figure 5-70. Nagatac lag-correlation: L₁ vs. QBO50

Nagatac (080005) Lag Correlations (T2 vs. QBO50)

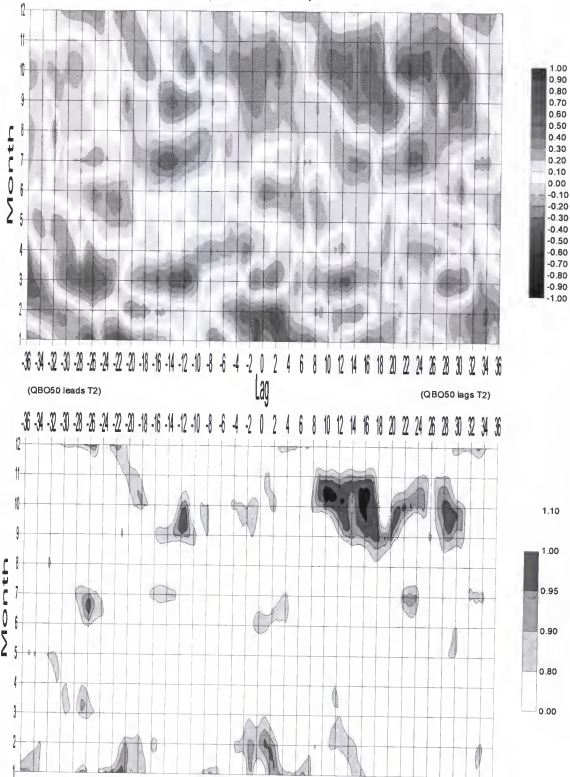


Figure 5-71. Nagatac lag-correlation: T₂ vs. QBO50

Nagatac (080005) Lag Correlations (T3 vs. QBO50)

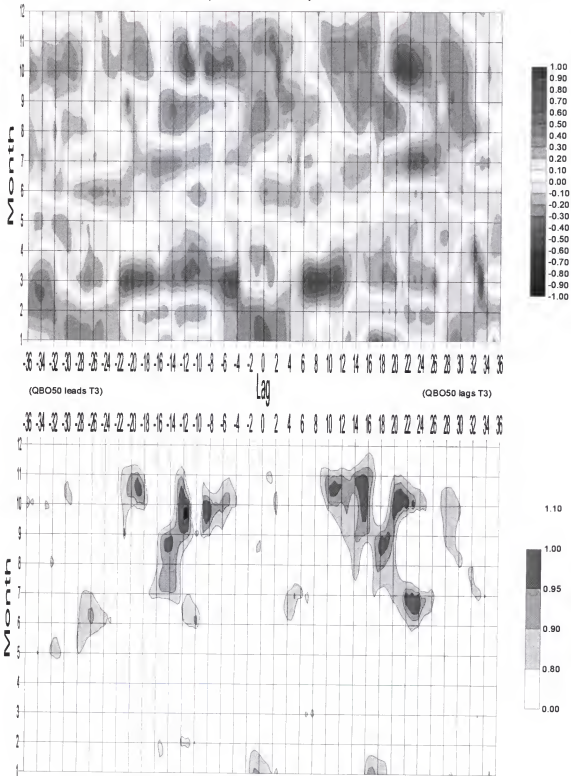


Figure 5-72. Nagatac lag-correlation: T₃ vs. QBO50

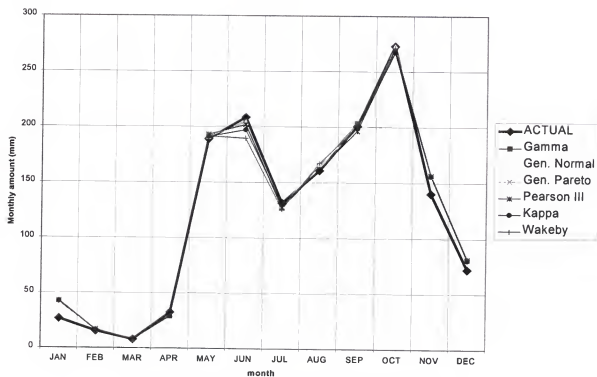
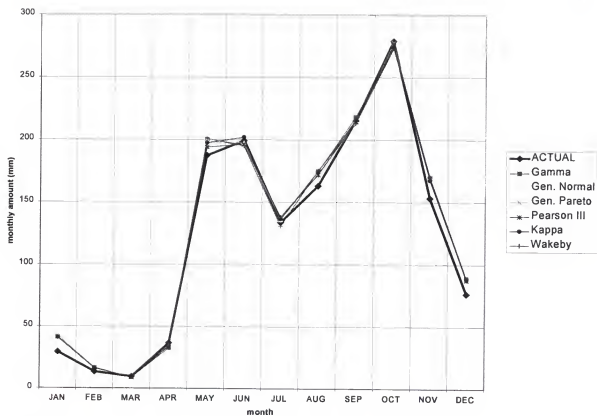


Figure 5-73. Rainfall simulation, Sanatorio Duran
All years (top) and "other" years (bottom)

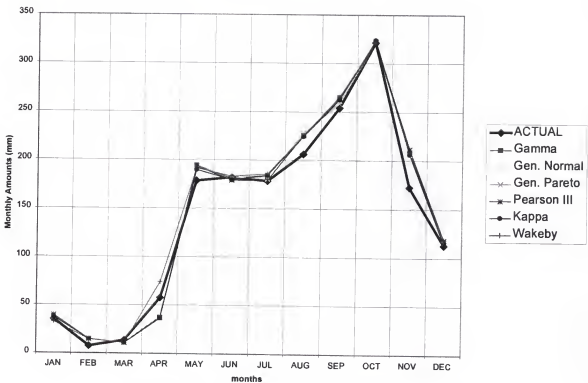
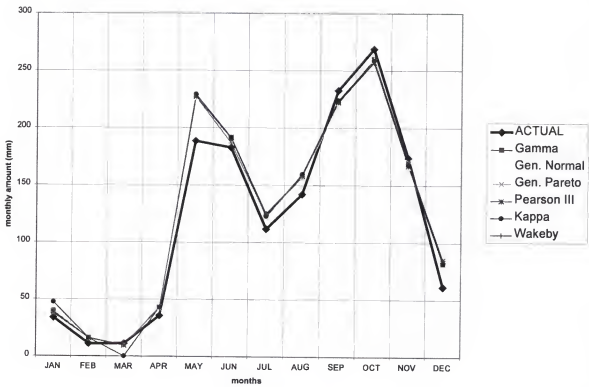


Figure 5-74. Rainfall simulation, Sanatorio Duran
El Niño years (top) and La Niña years (bottom)

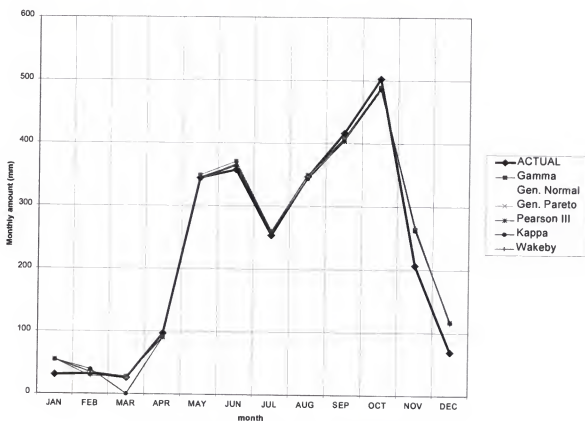
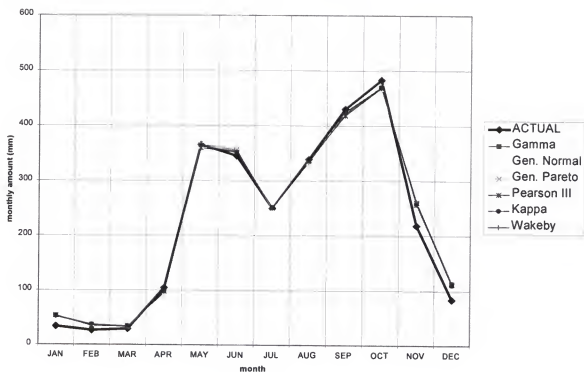


Figure 5-75. Rainfall simulation, Villa Mills
All years (top) and "other" years (bottom)

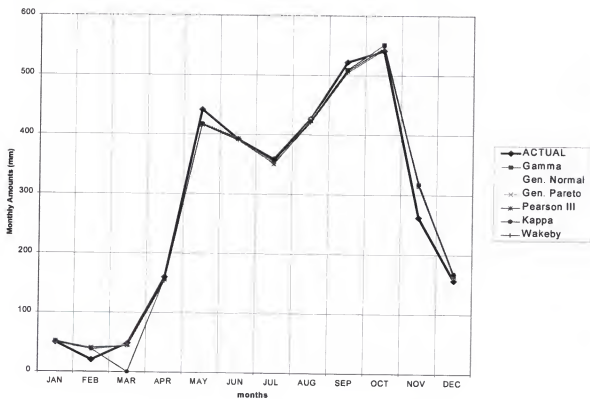
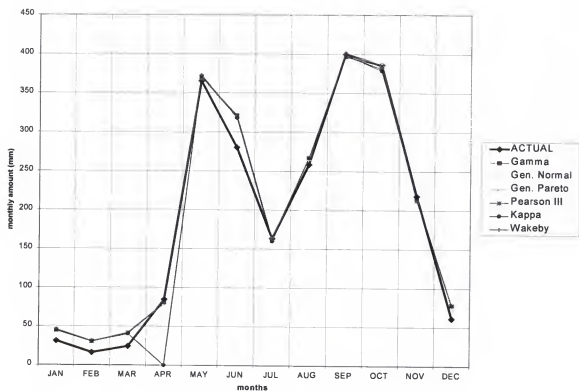


Figure 5-76. Rainfall simulation, Villa Mills
El Niño years (top) and La Niña years (bottom)

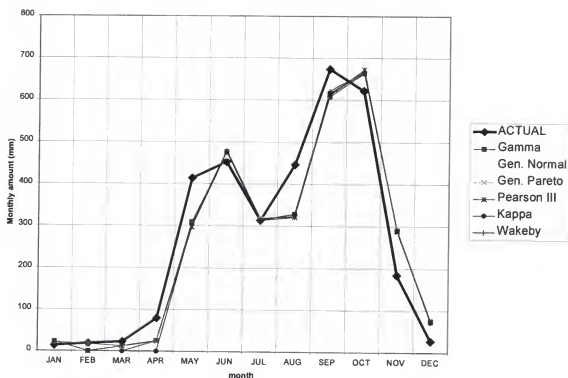
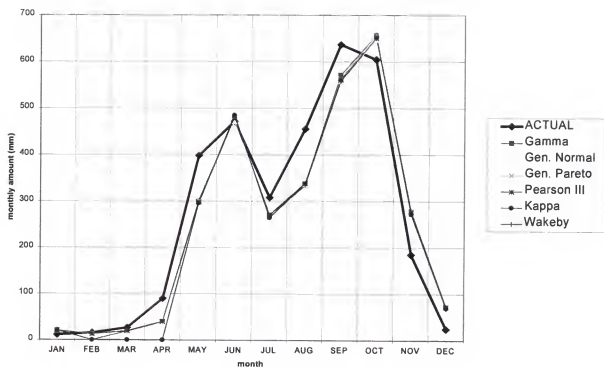


Figure 5-77. Rainfall simulation, Nagatac
All years (top) and "other" years (bottom)

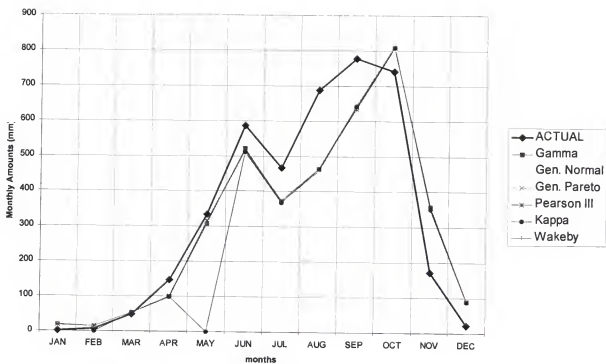
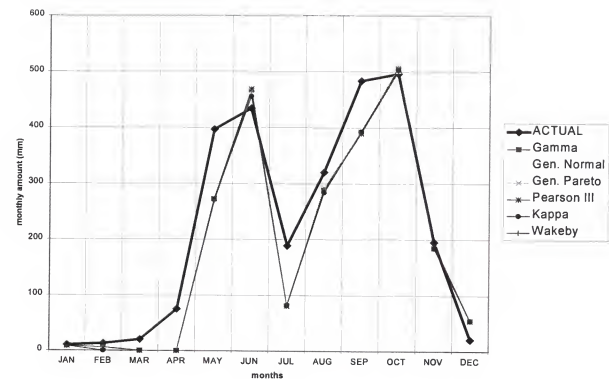


Figure 5-78. Rainfall simulation, Nagatac
 El Niño years (top) and La Niña years (bottom)

Sanatorio Duran (073011)

Bootstrap results -- January

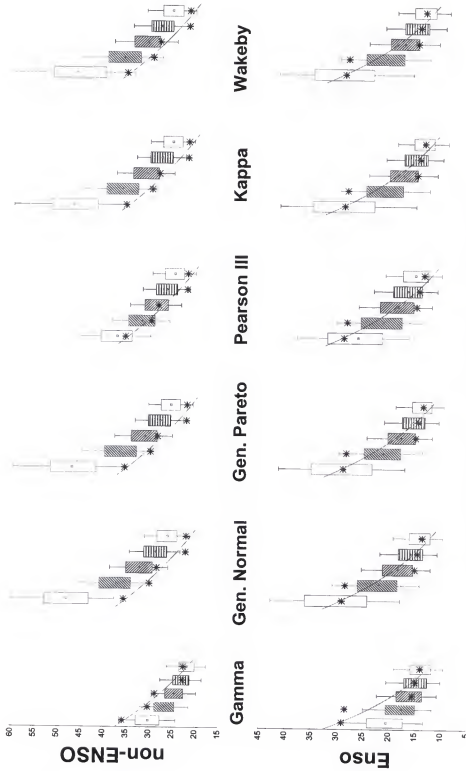


Figure 5-79. Sanatorio Duran bootstrap results, January.

Sanatorio Duran (073011)

Bootstrap results -- July

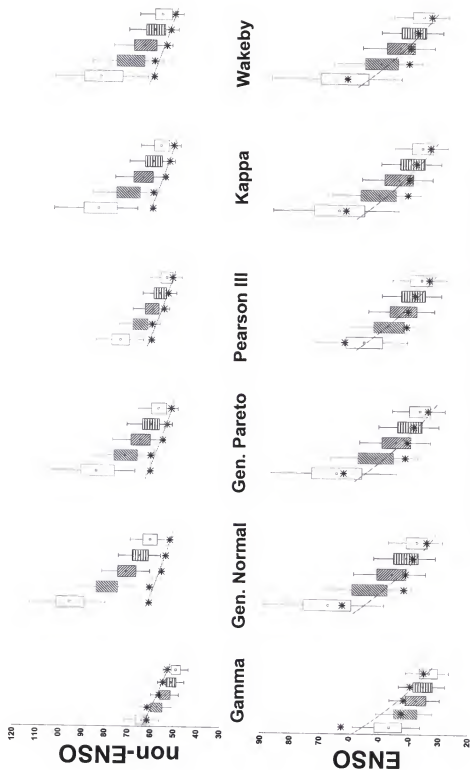


Figure 5-80. Sanatorio Duran bootstrap results, July.

Villa Mills (073033)

Bootstrap results -- January

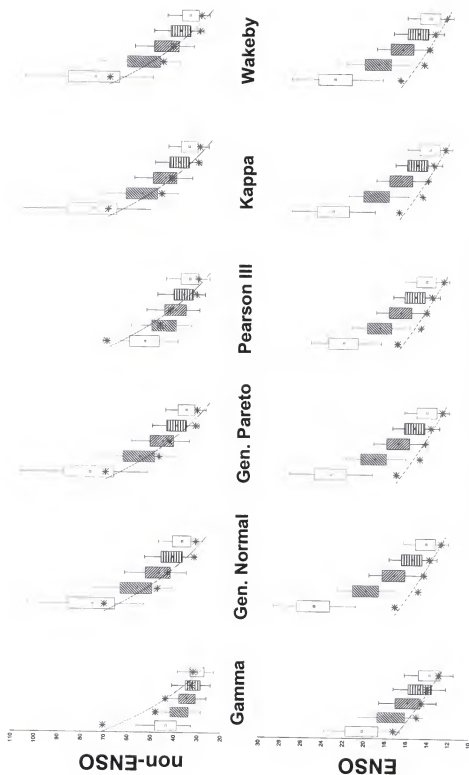


Figure 5-81. Villa Mills bootstrap results, January.

Villa Mills (073033)

Bootstrap results -- July

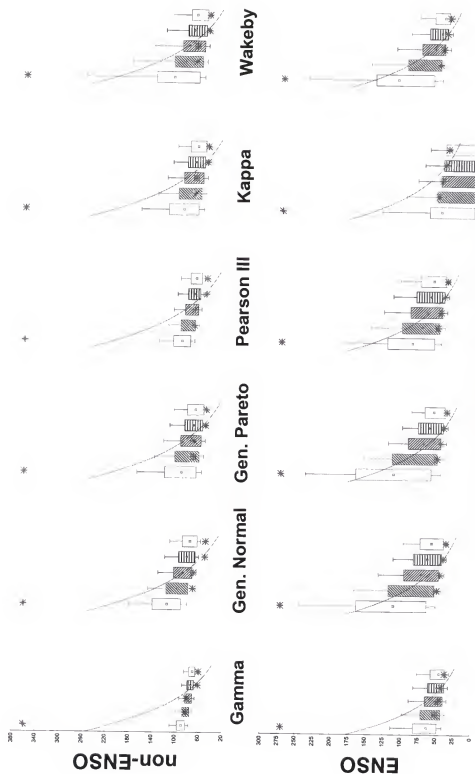


Figure 5-82. Villa Mills bootstrap results, July.

Nagatac (080005)

Bootstrap results -- January

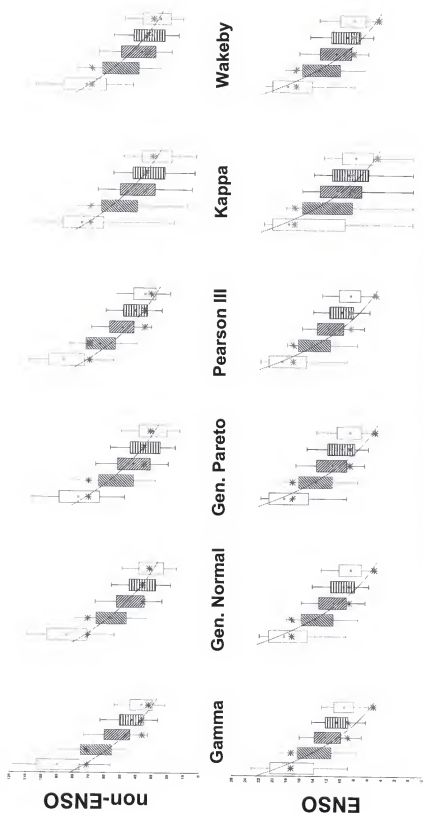


Figure 5-83. Nagatac bootstrap results, January.

Nagatac (080005) **Bootstrap results -- July**

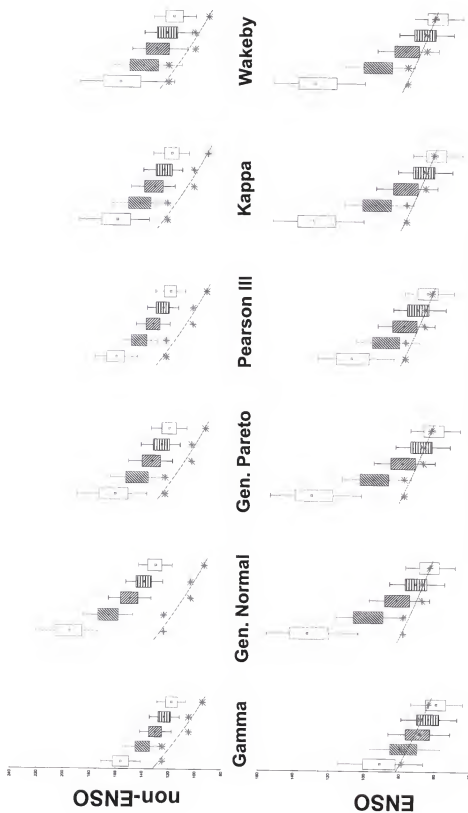


Figure 5-84. Nagatac bootstrap results, July.

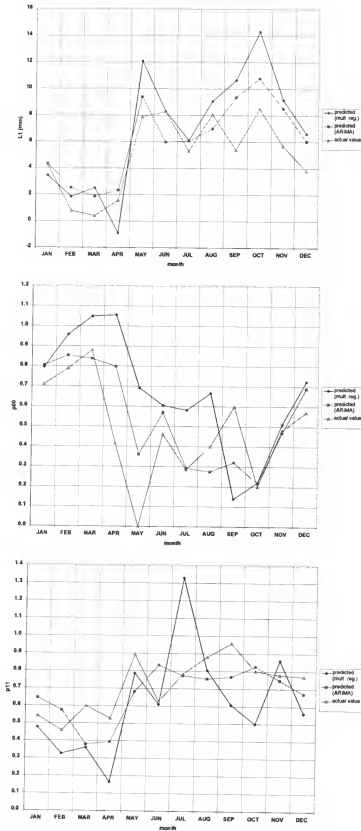


Figure 5-85. Sanatorio Duran parameter forecasts.

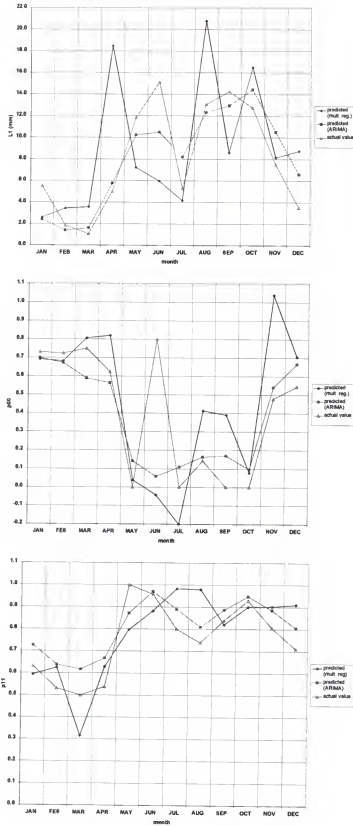


Figure 5-86. Villa Mills parameter forecasts.

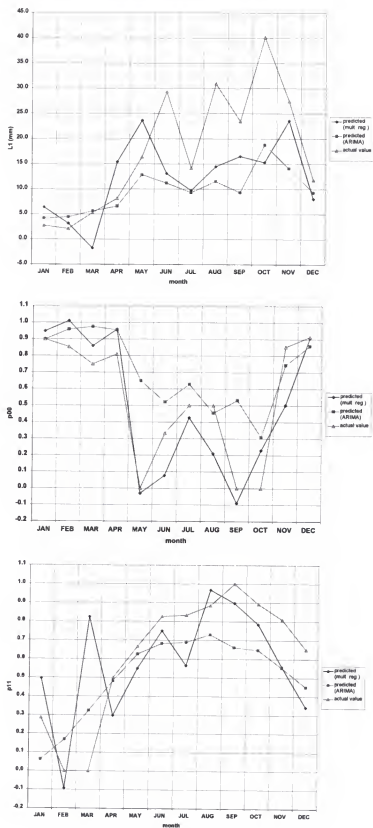


Figure 5-87. Nagatac parameter forecasts.

ARIMA results -- 073011

1991 forecast, January through December

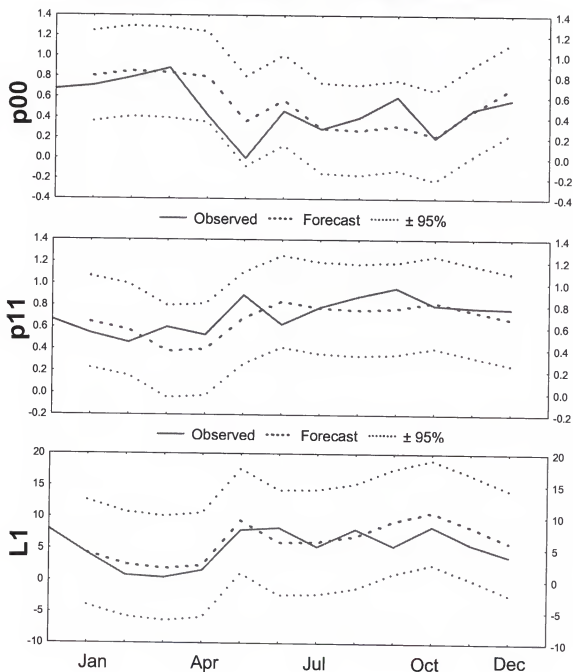


Figure 5-88. Sanatorio Duran ARIMA results, with 95% confidence interval.

ARIMA results -- 073033

1991 forecast, January through December

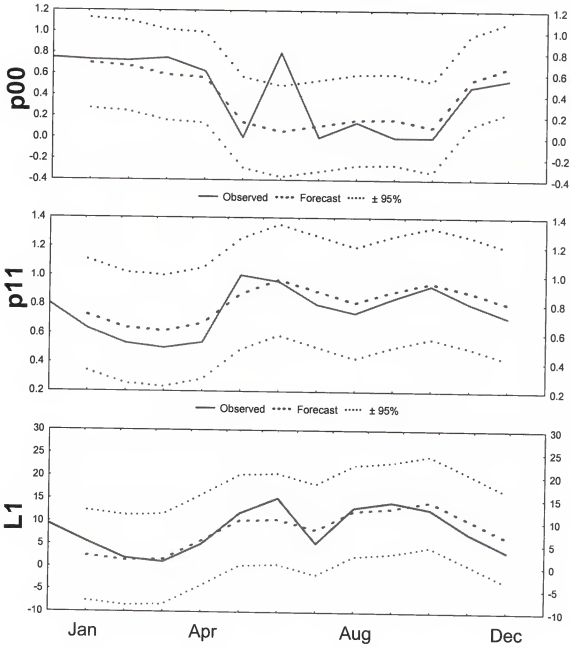


Figure 5-89. Villa Mills ARIMA results, with 95% confidence interval.

ARIMA results -- 080005

1995 forecast, January through December

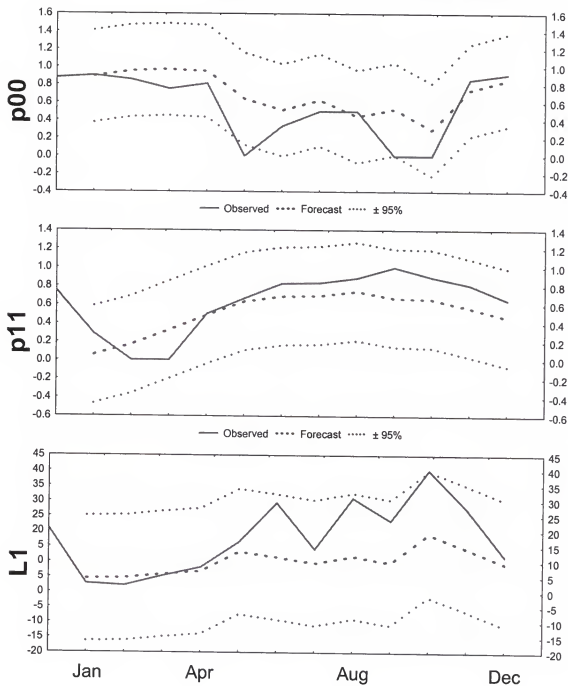


Figure 5-90. Nagatac ARIMA results, with 95% confidence interval.

CHAPTER 6

DISCUSSION

This chapter discusses the results noted in the previous chapter in terms of their overall, broad-scale meaning, both in terms of their geographic and atmospheric consequences. The results of the modeling exercise are then discussed in relation to the historical rainfall characteristics by way of a Monte Carlo simulation. Finally, the forecasting results will be discussed.

Model Identification

Markov Model Order

Both the AIC and BIC indicate that the zero-th and first order models predominate over higher-order models throughout the year. However, the AIC seems to be somewhat more biased towards the higher orders, and indicated by the relatively high numbers of high-order models selected compared to the BIC. This is in agreement with previous studies which found that the BIC tends to give more weight to economies of order as opposed to maximizing information contained within the model.

The distinction between predominating orders for summer and winter rains is especially important when examined in terms of the originating mechanisms of the rainfall during these seasons. It has been established (e.g., Waylen et al., 1994, 1996; Schultz et al., 1998) that the Costa Rican rainfall during the winter months can be attributed in large part to the passage of cold fronts, or *nortes*, which form over North

America and move southward across the pressure gradient that exists between North America and the Caribbean Sea. These fronts bring rainfall to Costa Rica in much the same fashion as they do in the temperate regions farther to the north. For that reason, Costa Rica may be thought of as acting like a temperate region rather than a tropical region, since the rainfall has the same originating mechanism and behavior. Previous studies (e.g., Katz, 1977; Stern and Coe, 1984) have examined rainfall in temperate regions, and concluded that the proper model is one of first order. Indeed, it is found that, during the winter months, Costa Rica should also be considered in the same fashion, based on the BIC results (Figures 5-1 to 5-3). Therefore, it can be concluded that, in terms of the statistical modeling of rainfall as a chain-dependent process, Costa Rica winters should be considered as temperate, rather than a tropical, climate.

The opposite holds true in the summer. Contrary to the indications of Katz (1977) and Stern and Coe (1984) indicate for temperate regions, Costa Rica rainfall occurrence should be considered as a zero-th order model. The reason for this relates to the sheer quantity of rainfall that the region receives during this period. Like most other humid tropical environments, Costa Rica experiences a distinct rainy season, during which convection-based rainfall occurs almost daily. Therefore, there is little reason to condition the rainfall probabilities on the previous day's occurrence, since it is a near-certainty that it rained on the previous day. Thus, the system holds little "memory" and the most efficient model is one which simply examines the long-term averages (i.e., "climatology") for the system, the zero-th order model.

The model-order conclusions are based solely on the information criteria (AIC/BIC) results, and thus may apply more importance to the parsimony of the model

than is warranted. Katz and Parlange (1998) have shown that lower order models, while more "efficient" from an information-criteria point of view, could lead to an "overdispersion" phenomenon which is attributable to the inability of a low-order model to account for high-frequency variation in the rainfall occurrence pattern. It is therefore important to understand the inherent limitations of choosing a low-order model when modeling rainfall occurrence. In addition, Jones and Thornton (1993) demonstrate that a third-order model does a n adequate job of simulating tropical rainfall in a tropical context.

The subtle shift in the zero-th order counts during the month of July is of interest, because it has been theorized (Hastenrath, 1990; Poveda and Mesa, 1997) that the *veranillos* are modified during El Niño years, due to alterations in the northward migration of the ITCZ during the summer, and an intensification of the Northeast Trades; in particular, the *veranillos* period tends to attenuate during non-El Niño years, while expanding during El Niño years. Since the data presented here are aggregated on a monthly basis, contraction of the *veranillos* would tend to eliminate the "bleed-over" into June or August, and thus concentrate its effects in July; on the other hand, during an expansion of the *veranillos*, such bleed-over would be more likely to occur, and its effects would be reflected in the June and August. When examining the BIC graphs for El Niño and non-El Niño years, such a pattern emerges. During non-El Niño July, there is a subtle reduction in the zero-th order counts and a commensurate elevation in the first order counts, indicating a greater likelihood that the rainfall on previous days is not as certain as the long-term average would suggest. El Niño year July, conversely, show very little change during the mid-summer months in the zero-th order counts. This supports the

view that warm ENSO episodes have a measurable effect on the duration of the *veranillos* period in Costa Rica.

The effects of ENSO on the *veranillos* illustrate why it is often better to use a “less efficient” (i.e., higher order model) (e.g., a first order over a zero-th order) rather than the model indicated by the information criteria. While a lower-order model might be more efficient, it can conceal important variation in the climatology of the region that can only be illuminated by considering the previous days’ conditions. It is also important to understand that higher-order models contain the same information that lower-order models carry, in addition to any additional information that is derived from conditioning on previous rainfalls, and for this reason, they can be more useful than the more efficient lower-order models. In the case of the present study, we wish to understand not only the occurrence of rainfall on a particular day, but how that rainfall occurrence is altered on a day-by-day basis by the state of the ENSO. For this reason, it is justifiable to use a first order model when examining the characteristics of daily rainfall occurrence in Costa Rica, even though such a model may not be the most efficient system. However, for purely practical applications (e.g., the creation of agricultural and hydrological predictive models), the most effective model for summer rainfalls would seem to be of zero-th order.

First Order Markov Model Results – Segregated Years

The presence of numerous climatological forcing mechanisms in league with the extreme terrain variability inherent in the Costa Rican region, combine to form a very complex temporal pattern for the El Niño versus non-El Niño years across the study transect. These effects are less pronounced in winter than in summer, due to the

decreased overall precipitation in the region. It is, however, important to note that the effects are more apparent in that part of the region which is more directly exposed to the *nortes* (Sanatorio Duran and Villa Mills), rather than in the extreme west, as represented by Nagatac, where very little precipitation occurs during the boreal winter. The lack of seasonal rainfall in the west is reflected in the winter values of p_{00} and p_{11} , which are respectively very high and low; this remains true regardless of the year classification, and is reflective of the extreme aridity which dominates in the winter. This can be accounted for by the southward migration of the ITCZ during November and December, leading to a resumption of Trade wind domination, and the associated rainshadow effect, across the region during the winter months. This high pressure scenario is aided by the persistence of the Costa Rica Dome, which was intensified during the preceding summer months by the Trade winds across the area (Umatani and Yamagata, 1991).

An important winter pattern develops in the eastern part of the transect, where the *nortes* are more strongly felt. The increase in the winter p_{11} (and commensurate decrease in p_{00}) for La Niña years for both Sanatorio Duran and Villa Mills can be explained in terms of the relationship between the *nortes* and the ENSO conditions prevailing during that year. The supposition by Waylen et al. (1994, 1996) and Schultz et al. (1997, 1998) that *nortes* are inhibited during El Niño years and enhanced during La Niña years is supported by the data. The effect is most striking at Villa Mills, possibly due to the greater northward exposure of the station compared to Sanatorio Duran, which is located on the southern slope of Volcán Irazú, and is therefore less likely to receive the full benefit of cold surges originating to the north. Villa Mills, moreover, located at a greater altitude than Sanatorio Duran as well as being farther south, which could account for the

greater degree of separation between year types; La Niña years seem to be associated with stronger and more frequent *nortes* early in the winter season (November through January), and the increased intensity could allow for more recurrent rainfall at higher elevations, and at more southerly latitudes. Thus, it would rain more often at Villa Mills during the winter for La Niña years compared to other year types. Conversely, El Niño years would have the same sort of effect in an opposing fashion, being more prone to hinder rainfall frequency at Villa Mills than at Sanatorio Duran, increasing p_{00} and decreasing p_{11} .

The analysis of the summer moths is somewhat easier, due to the overwhelming strength of the two dominating rain-inducing mechanisms: the increased Trade winds and the northward ITCZ migration. During early summer (May and June), the increased Trades bring much higher rainfall totals to the eastern part of the study area. The nearly uniform p_{11} at Sanatorio Duran and Villa Mills during this period seems to suggest that the length of rainy spells in the east and central part of the transect is relatively unaffected by the state of the ENSO. However, there is distinct decrease in the p_{00} parameter during the same period for La Niña years (although there is little or no change in the El Niño p_{00}); it can therefore be interpreted that, while El Niño years do not seem to affect early summer rainfall frequency, La Niña years are more likely to experience fewer consecutive dry days between multi-day rain events. The best physical explanation for this is that, while Sanatorio Duran and Villa Mills are located in the Reventazón basin, both stations continue to receive Pacific-type influences due to "bleed-over" of ITCZ-induced storms across the cordillera; this is especially true of Villa Mills, which is located in close proximity to the crest of the cordillera. (See George et al. (1998, Figure

6) for a detailed illustration of this process.) The overall effect, when viewed from the standpoint of rainfall frequency alone, is to have more rainfall events occurring at Villa Mills and Sanatorio Duran during early summer La Niña years than during El Niño years.

Mid-summer (July and August) represents the timeframe of the *veranillos* in Costa Rica. The primary explanation for this phenomenon is the reestablishment of the Trade winds across the isthmus, which helps to restrain the northward migration of the ITCZ. This causes the sudden decrease in p_{11} and commensurate increase in p_{00} across the study area. The effect is most pronounced in the west, with Nagatac clearly demonstrating the decrease in rainfall frequency in July, as evidenced by the decrease in p_{11} and increase in p_{00} . This effect is much more pronounced during El Niño years, which is clearly seen in Figure 5-4 through 5-6, in which El Niño years have a substantially lower p_{11} and higher p_{00} , reflective of the sensitivity of western Costa Rica to changes in the state of the ENSO.

The Nagatac p_{xx} graphs (Figure 5-6) also demonstrate the extraordinary influence of the western equatorial Pacific conditions on the migration characteristics of the ITCZ in the eastern Pacific; the rainfall at Nagatac could in this instance be viewed as a proxy for the ITCZ position and strength. Clearly the state of the ENSO influences the ITCZ, such that during El Niño events, eastern Pacific ITCZ northward migration is inhibited. The limiting factor is the strength of the Trade winds, which increase as warm waters are advected into the eastern Pacific region. The increased winds offshore of western Costa Rica also help to enhance the upwelling of the ocean waters, expanding the Costa Rican Dome, and helping to stabilize the atmosphere, which in turn further inhibits northward migration of the ITCZ. Thus, the increased Trades act to inhibit western Costa Rica

rainfall in several ways: by “pushing” the ITCZ-related storms southward, and by creating an unfavorable atmospheric environment for further storm development.

The effect is less pronounced in the eastern part of the transect in terms of El Niño versus La Niña differences during the *veranillos* proper (July/August). However, there is an intriguing trend which develops for Sanatorio Duran and Villa Mills during the early and mid summer periods, especially when examining p_{00} . While there is little change in p_{00} between El Niño years and “other” years, La Niña years consistently have a lower p_{00} throughout summer. This would seem to indicate that La Niña years would have moister summers overall (when considering rainfall frequencies alone) than other years, since a lower p_{00} would argue for less dry-day spacing between rainfall events. Waylen et al. (1994, 1996), however, have shown that, on balance, the eastern part of Costa Rica is moister overall during El Niño years; since La Niña years tend to nearly always occur either immediately prior or subsequent to El Niño years, these results would appear, on the surface, to be contradictory to existing observations. For this reason, it is important to examine the frequency modeling within the context of magnitude models as well, rather than as a stand-alone result. One possible physical interpretation is that these stations are receiving the same type of “bleed-over” previously discussed. During the La Niña years, the ITCZ intensifies and moves farther to the north compared to El Niño years; increased storm activity to the west could reach sufficient strength to allow penetration of the cordillera, allowing more westward-originating precipitation to reach deeper into the Central Valley than during El Niño years. This process would be enhanced by the relative decrease in Trade wind strength during La Niña years. The observed late-summer (September-October) increase in p_{11} , and concurrent reduction in p_{00} , for La Niña years,

most pronounced at Sanatorio Duran and Villa Mills, is a continuation of this same process. Ultimately, it is difficult to make firm statements about the behavior of the entire Caribbean fringe region from these two stations, because neither station, while best representing the most easterly part of the study area, is entirely outside of the Central Valley, and therefore is subject to a wider range of climatic influences than stations located farther to the east. The continuation of this research using data from more easterly stations would represent an important addition to the complete understanding of the climatic nature of this region.

L-Moment Model Results – Segregated Years

Sanatorio Duran. Sanatorio Duran shows a pattern that seems to contradict its easterly location, both in terms of its total monthly amounts and the L-moment parameters. Stations located on the east side of the isthmus will receive higher total precipitation amounts during ENSO events, due to the increased strength of the Trades during this period. However, Sanatorio Duran seems to act more like a station with “Pacific” characteristics, in that the La Niña years give higher precipitation totals for the entire summer, as well as higher L_1 values for the mid- to late summer. The likely explanation hinges on its geographic location with respect to Volcán Irazú; its southern orientation would not only shield it from many of the Caribbean processes, but also align it towards any Central Valley storms resulting from Pacific-based processes.

It is important to note that there are two periods in which Sanatorio Duran displays “Caribbean” attributes. The first is early Summer, prior to the *veranillos*, during which the L_1 parameter is substantially higher during El Niño years than during La Niña years. This reflects the fact that the Trades dominate the entire eastern half of the country

while the ITCZ is still fairly far south. During this period, the storms which originate in the east intensify, producing more rainfall in El Niño years, and resulting in the elevated L_1 statistic; on the other hand, the relative increase in p_{00} seems to suggest that, while the storms get bigger, there is a greater likelihood that they will be spaced farther apart, resulting approximately the same amount of rainfall occurring during the early summer, regardless of the type of year.

The second time period in which the rainfall is more characteristically “Caribbean” is during the winter, specifically the late winter (February/March). Overall, the lack of significant change in p_{00} throughout the winter months indicates that there is little change in the frequency of cold air masses arriving from North America; in other words, the frequency of cold front arrival, at least as far as Sanatorio Duran is concerned, is unaffected by the state of ENSO. However, during the late winter, the duration of the individual storms (indicated by p_{11}) and the relative strength of the storms (measured by L_1) are measurably different for different year types. During El Niño years, the duration of the individual storms seems to increase somewhat, without any change in the amount of daily rainfall produced; during La Niña years, the strength of the storms seems to increase in late winter, without much change in the duration of the storms. This is an important observation, because it reflects the observed general tendency of the *nortes* to be depressed during ENSO events.

Villa Mills. Villa Mills seems to possess climatic properties of both the eastern and western part of study area. During the winter, it shows evidence of the *nortes*, in much the same fashion as Sanatorio Duran; a slight increase in daily rainfall amounts (Figure 5-9) during La Niña years above what occurs during El Niño years is seen,

indicating that *norte*-related storm intensity is attenuated during ENSO events. During the summer, Villa Mills shows Pacific characteristics; La Niña daily rainfall amounts (as evidenced by L_1) are uniformly higher, and El Niño years show much lower daily storm amounts. A plausible explanation for this derives from Villa Mills' unique geographic location: it stands at a relatively high elevation (approximately 3000 meters), and is quite close to the divide separating the Caribbean and Pacific drainage areas. Thus, Villa Mills is the recipient of climate processes from both sides of the Costa Rican isthmus, and will behave as a "Caribbean" or "Pacific" station, depending on which set of processes is dominant. During El Niño years, while the Trades dominate farther east, there is considerable "bleed-over" from the Pacific effects, reducing the precipitation during El Niño events and enhancing the rainfall during La Niña years. The overall domination by Pacific influences during the summer is probably linked to the orientation of the intervening *cordillera* as seen when looking east and west of the station; the *Cordillera Central* acts as a blocking influence for the Northeast Trade winds, reducing the effectiveness of the winds for advecting moist air into the region during the summer, and allowing the station to behave as if it were located farther west. For this reason, Villa Mills rainfall are less frequent, and less intense, during El Niño years

Nagatac. The p_{xx} and L_1 results for Nagatac seem to indicate that the station acts in a "classic" Pacific fashion, in that the summer shows a definite *veranillos* pattern during July; this corresponds to the period in which the Trade winds strengthen and inhibit the northward ITCZ migration. Also shown in Figures 5-6 and 5-7 is the increase in overall precipitation in the late summer, with L_1 reaching a maximum in October. As

previously noted, however, p_{11} reaches its maximum (and p_{00} its minimum) about a month earlier for all year types except El Niño years.

The best physical explanation for this apparent contradiction is that the influence of the ITCZ increases gradually over the course of two to three months. As it migrates northward, Nagatac begins to experience increased rainfall frequencies (i.e., increased p_{11}) as the northern fringe of the ITCZ encroaches on the region. These storms would build in strength over several months as the core of the ITCZ moves farther north. In this way, the region would initially experience an increase in storm frequency (increased p_{11}) and later an increase in the magnitude of the resulting storms (increased L_1). This pattern is observed in the Nagatac record.

The winter results show an interesting pattern, in that while the total monthly precipitation is extremely low (Figure 3-3), the rainfall magnitudes (L_1) do not seem to be reduced by a commensurate amount; L_1 readings are about half of the average summer values, while winter total monthly averages represent about 1% of the summer values. Clearly, the bulk of the change from summer to winter precipitation is accounted for in the frequency of the rainfall, represented by the p_{xx} findings (Figure 5-6). Physically, this is the result of the absence of rainfall-inducing mechanisms in the region; the ITCZ is at its southernmost position during this period, the Trade winds are relatively light, and the offshore waters cooler than during the summer. The only significant rainfall-producing process is the occurrence of the *nortes*, which affect primarily the northeastern part of the country. It is clear that the amount of "bleed-over" from the east of *norte*-based rainfall is at least partially dependent on the state of the ENSO; during El Niño years, the magnitude of the resulting rains (L_1 , Figure 5-11) is smaller than during other year types.

The is consistent with other findings (e.g., Schultz et al., 1997, 1998), in which the cold air mass intrusion from North America is attenuated during El Niño years.

Time Series Correlation – L-Moments and Transition Probabilities

Performing a lag-correlation on the “pure” (i.e., unsegregated) data time series allows us to avoid the somewhat artificial designations of “El Niño” and “La Niña” years. The lag-correlated findings are, therefore, the integrated results of all of the various atmospheric processes which contribute to the generation of rainfall, as they relate to a single ocean-atmosphere index. As discussed in “Methodology,” the indices chosen for this study are those which are most likely to contribute to precipitation in Costa Rica. An examination of Tables 5-13 to 5-71, which summarize the overall correlations between the rainfall parameters(p_{xx} , L_1 , T_2 , T_3) and the ocean-atmosphere indices (SOI, NAO, NINO34, QBO50) show that, for the most part, the response types can be partitioned into “sub-seasons” (e.g., early summer, late summer, etc) which correspond to well-understood patterns of monthly rainfall. The winter and summer seasons are clearly distinguished.

In general, the interpretation of the p_{xx} and L_1 figures is relatively straightforward: a strong positive L_1 correlation indicates that, as the index rises for a given lag (with a negative lag indicating that the index value was recorded a certain number of months prior to the measured statistic), the statistic would rise as well. Thus, a strong positive correlation between the SOI for lag = -2 months and p_{11} in January would mean that the January p_{11} rose when the SOI for the November immediately prior fell. In this way, the p_{00} , p_{11} , and L_1 can be immediately interpreted as measures of sustained rainfall, sustained dry spells, and rainfall intensity.

However, interpreting T_2 and T_3 is more difficult. T_2 is a measure of the degree of variation in the rainfall intensity, while T_3 measures the skewness of the rainfall intensities. It is possible to experience an increase of one or both of these statistics in relation to a given index. Since rainfall intensity has a fixed lower limit (in the case of the present study, this limit is the minimum measurable rainfall, 0.1 mm), any variation and skew must be the result of higher and/or lower values towards the higher end of the distribution. In general, if the correlation between an index and the T_2 statistic is elevated in either a positive or negative direction without a commensurate change in T_3 , the conclusion must be that rainfall intensities have not been changed to any significant degree, but rather that the highest rainfall amounts remain unchanged in value while more of these same rainfall amounts have fallen. Conversely, if the T_3 correlation shows considerable change without a similar change in the T_2 measure, then the rainfall maximums must have been changed without very many of the large events occurring. It is within this context that the T_2 and T_3 lag-correlation diagrams must be interpreted.

When examining the diagrams, it is also important to recognize that different processes are in control for the summer versus the winter seasons. This becomes evident when viewing the lag-correlation. There are several possible explanations for this. The most obvious has to do with the average frequency of rainfall over long time periods for summer versus winter; during summer, it is not unusual for many locations to receive rainfall nearly every day, regardless of the state of ENSO, and for this reason it could be that changes in the ENSO might result in a change in rainfall amounts, rather than a complete cessation.

A second explanation involves the degree to which the "standard" indices can measure the impact of ENSO on Central America. This is especially relevant because the greatest rainfall amounts occur during the boreal summer, yet the maximum extent of the ENSO occurs during the boreal winter. Thus the period of maximum rainfall in Costa Rica is out of phase with the period of warmest waters in the eastern equatorial Pacific Ocean. It is important in this context to understand the relationship between the increased ocean temperatures in this region (which serve to promote the development of the ITCZ), and the Trade winds which advect most of the humid Caribbean air across the isthmus and bring rains to the eastern part of the country; the equatorial Pacific SSTs are well-represented by the NINO34 index, but to date no suitable long-term index has been developed to summarize the influence of the Trades. Coupled with this is the fact that the initial forcing mechanisms which motivate the onset of an ENSO event are far from being fully understood. Therefore, it is not unlikely that the ability to use index values from six months to a year prior to a period of interest would provide less than perfect insight into the day-to-day transition of rainfall across the region. When the events become more synchronized, as they do during the boreal winter, the dependence on ocean-atmosphere conditions becomes more acute. None of this is to be interpreted as being contrary to previous findings of summer ENSO/precipitation relationships; the present research was designed to reduce the rainfall models in the region to their smallest elements, and for that reason the results cannot be viewed independently.

Overall, the roll of the SOI is certainly less impressive than what might have been previously expected; the NINO34 index, which provides a direct gauge of Pacific warming during an ENSO event, appears to have stronger associations with the lag-

correlation of rainfall in Costa Rica. Nevertheless, while none of the indices seem to present themselves as outstanding predictors of rainfall characteristics, this simply confirms that the climatological picture of Costa Rica is far more complex a system than can be simply modeled using a small set of Pacific Ocean data.

Winter season rainfall correlations. The p_{00} and p_{11} parameters are strongly affected at Sanatorio Duran and Villa Mills during December and January. This effect is much more evident for the NINO34 correlations than with the SOI. Also, Sanatorio Duran and Villa Mills seem to display a higher degree of correlation throughout the winter. The conclusion that can be drawn is that the *norte* activity which is responsible for the bulk of the rainfall activity is clearly associated with the temperature of the equatorial Pacific waters. The negative correlations between NINO34 and p_{11} for the months prior to the winter rainy season, and the positive correlations with p_{00} during the same period, lead to the conclusion that the *nortes* are inhibited during an ENSO event. It is interesting to note that the L_1 parameter shows a positive correlation to the NINO34 indices for the months just preceding the winter season for Sanatorio Duran and Villa Mills. When viewed along with the p_{xx} results, this would seem to suggest that while the frequency of rain events decreases during an El Niño, the actual rainfalls that occur are of higher magnitude than other years.

For the most part, Nagatac does not display many significant correlations for the winter months. This is probably due to the extremely dry conditions that prevail in the western areas of Costa Rica during this period. One important exception is the p_{11} /NINO34 correlation during March and April. The strong positive correlation at lag = -12 months, as well as the negative association at lag = +12 months, is suggestive of a

linkage between the *norte* activity and the central Pacific SST. The correlations would seem to indicate that the length of rainy spells decrease during the late winter in the year preceding the height of an El Niño event, while they increase in duration in the year following that same event. This inference is supported by Schultz et al. (1997, 1998), in which *norte* activity is modulated by El Niño occurrence. The lack of rainy days during March in Nagatac compared to the eastern portion of the study area would make the p_{11} statistic very sensitive to high-frequency changes, such as those associated with ENSO.

The fact that most of the strongest associations occur in winter for the eastern stations seems to indicate the degree to which each of the stations is susceptible to changes in the occurrence of *nortes*. Each station shows a significant negative p_{00} , and positive p_{11} , correlations with the SOI during December and January. This provides a considerable explanation for the decrease in rainfall over Costa Rica during the winter during ENSO events. This response is further reinforced by the NINO34 winter correlations, where the p_{00} tends to be positively correlated and the p_{11} is negatively correlated. These findings reinforce the idea that the frequency of *nortes* are significantly influenced by the state of the ENSO; during El Niño events, *nortes* tend to be significantly reduced. Somewhat less in evidence is the “flip-flop” nature of the *nortes* response in relation to the timing of ENSO events, especially at Sanatorio Duran; the p_{xx} /NINO34 correlation diagram (Figure 5-15) clearly show a reversal in significant correlation values for December/January, which are spaced approximately one year apart. This helps to confirm that the *nortes* experience a resurgence either the year prior to, or following, an ENSO event.

It has been suggested (Schultz et al., 1997, 1998) that the alterations in hemispheric-scale air flow during an El Niño event could explain the decrease in *norte* occurrences; the cold air masses which are able to overcome the atmospheric blocking effects associated with changes in the subtropical jet stream could be strong enough to enhance precipitation magnitudes in Costa Rica once the cold fronts arrive in that region. This process could account for the increase during El Niño episodes in L_1 and p_{00} , while manifesting a decrease in p_{11} .

Summer season rainfall correlations. The summer correlations for the three stations are interesting on several fronts. First, there is a lack of significant correlations during much of the summer season (e.g., early summer p_{00} , early summer L_1); this indicates that at least some of the important driving mechanisms in the Trades/ITCZ scene are being missed with the present indices. As mentioned before, there are few ways of deriving a reasonable "Trade Wind Index" since such measurements have not been accurately recorded for extended time periods. The ITCZ, on the other hand, is an easily definable entity, since its presence is noted by extensive cloud cover and warm waters. Thus, the use of the standard indices allows only a partial view into the changes the Trade winds undergo during El Niño

One important observation that can be gleaned from Tables 5-1 to 5-3 is that, for the most part, the significant correlations between L_1 and the indices are confined to Sanatorio Duran, while the majority of p_{xx} -related correlations involve all three stations. In addition, Nagatac seems to experience very little in the way of L_1 correlations during any season. This could be interpreted to indicate that the ITCZ is itself not greatly affected in strength by the state of the ENSO, at least not in the vicinity of Costa Rica;

rather, its position is affected, and the direct result of the southward suppression of the ITCZ during El Niño events manifests itself as changes in the number of storms being produced.

Another important feature of the summer season involves the mid-summer months, the period associated with the *veranillos*. Comparisons can be drawn between Sanatorio Duran and Nagatac, since they lie on opposite sides of the study transect. Sanatorio Duran shows a strong synchronous negative correlation between L_1 and NINO34 (with positive T_2 and T_3 correlations), and positive correlation to the SOI coupled with a negative correlation for T_3 ; this indicates that the magnitudes of rain events is suppressed during ENSO events on the Caribbean slope of the study area. During the same time period, Nagatac shows a positive p_{00} association with NINO34, and negative association with the SOI. Similar associations appear when examining the L -moments; the SOI/ L_1 correlations are positive in July, while the NINO12/ L_1 correlations are negative. This demonstrates the sensitivity of the Pacific slope to changes in the *veranillos*, which are in turn linked to the state of the ENSO. Thus, it can be shown that while the net result for both sides of the country is a decrease in precipitation during the *veranillos* for El Niño years, the reasons for this decrease in the east are less intense storms, and the west experiences fewer storms of overall lower intensity.

The interpretation of the QBO50 diagrams are somewhat more problematic, because the physics surrounding the connection between long-term climate change and troposphere/stratosphere interaction are poorly understood at present. To complicate the interpretation, there has been a demonstrated statistical linkage between the behavior of the QBO and low-atmosphere climate anomalies. The few studies which have been made

on the subject concentrate primarily on statistical linkages (Hamilton, 1990; van Loon and Labitzke, 1990; Kodera, 1991; Salby and Shea, 1991; Mason and Tyson, 1992; Labitzke and van Loon, 1993; Kazadi and Kaoru, 1996; Tinsley, 1997), since little is known about the physical linkages between tropospheric/stratospheric interaction and climate anomalies.

Sanatorio Duran shows an significant relationship between p_{xx} and the QBO50 index. Clearly evident is the quasi-biennial cycle, especially for p_{11} during October/November. The same quasi-biennial cycle appears in April for p_{00} , though the correlations are somewhat less significant. It is important to note two things about the correlation signature at Sanatorio Duran. First, the April and October/November cycles, while having similar periods, are substantially out of phase; this suggests little association between April p_{00} and October p_{11} parameters, that is, the two parameter correlations are independent. Second, the timing of significant correlations (April and October/November) correspond to periods in which the transition from summer to winter (and vice versa) characteristics is occurring. This suggests that the effects of changes in the QBO50 index would manifest themselves either during the time in which neither tropical nor extra-tropical mechanisms are dominant (as in March/April), or when the process is nearly homogenous and dominated by a single mechanism (as in October/November, when the ITCZ overrides all other processes). If this is true, it could indicate that the effects of the stratospheric winds is of minimal importance from a physical standpoint, and that its effect would only become apparent when other mechanisms are in abeyance. However, during these periods, the role of the QBO50 as a

forecast tool cannot be overlooked; Figures 5-70 shows that the relationship, while not clearly understood, is certainly real and significant.

It is also notable that this effect is confined to Sanatorio Duran for the p_{xx} statistics, but dominates for L_1 at Nagatac. As mentioned before, the physical connection between the QBO and rainfall generation is poorly understood at present, but the idea that the QBO would primarily affect the frequency of daily rainfall in the east while affecting the magnitudes of daily rainfall in the west is intriguing. The only way of further examining this area is to expand the study area eastward and test whether the pattern holds. This represents a fruitful area of future research.

Model Choice, Validation, and Rainfall Simulation

Initial Model Choice. The use of a Chi-squared test to make the initial choices of candidate distributions is a contentious issue. It is well known that a necessary precursor to performing the test is the categorization of the data into “bins”; if any of the bins fails to have a sufficient number of members (usually five), then the test results are inconclusive (Norman and Streiner, 1986, pg. 81). The use of several decades of daily data binned into five groups, as was done in this study, insures that all bins will contain sufficient data points. Certainly, the test cannot be used to make a final determination as to distribution type, but it is adequate to eliminate the most inappropriate models from consideration.

An examination of Table 5-4 demonstrates that the Chi-squared test is effective in eliminating those distributions which lack parameters that allow an adjustment in the skew. The two-parameter Normal, Gumbel, and Exponential distributions are limited in that the amount of skew present in the distribution is fixed, and therefore the right-hand

tail of the model cannot be adjusted to accommodate extreme values while still effectively modeling the smaller and intermediate values which show up more frequently.

The exclusion of the two three-parameter distributions as a result of the Chi-squared test (the Generalized Extreme Value (GEV) and the Generalized Logistic (GLO)), most likely results from their specialized nature. The GLO is a "growth curve" whose basic application is the asymptotic approach to a known maximum value (Shaw and Wheeler, 1985). The distribution is therefore self-limiting on the right-hand tail, and inadequate for describing distributions with extremely large values compared to the minimum values which form the bulk of the data. The GEV, on the other hand, was explicitly constructed to describe the characteristics of extreme values of other distributions (Haan, 1977), and generally lacks the flexibility to fit both the upper and lower limits of a raw daily data stream.

The models which survive the test have some common characteristics. Most important, they all possess a "shape parameter" which allows the skew of the resulting distribution curve to be adjusted. This allows the extreme values to be better modeled than the Normal, Gumbel, and Exponential curves. Also, several of the remaining curves are mathematically related to one-another. The Pearson III distribution is a generalization of the Gamma, with the addition of a "location parameter" which in general should allow the Pearson III curve to better fit the extreme data values. Also, the Wakeby model is basically a double-Pareto mixture model, which would indicate that the Generalized Pareto and Wakeby models would act in somewhat similar fashions when further examination is complete.

Rainfall simulations. The rainfall simulations imply that it makes little difference which distribution is chosen among those remaining after the Chi-squared test. Indeed, Figures 5-73 to 5-78 show an exceedingly close correspondence between the average monthly rainfall amounts and the simulated monthly amounts. There are several conclusions which might be drawn from this.

First, it would seem that, from the standpoint of monthly rainfall, magnitude distributional choice is less important than the choice of rainfall occurrence parameters, p_{00} and p_{11} . This is logical since the choice of whether rainfall has occurred on a given day is made prior to the choice of a rainfall amount, and the decision whether to calculate a rainfall depends entirely upon whether it rains at all on a given day. It would seem, therefore, that the choice of distribution is more critical during those months when the p_{11} is relatively high and p_{00} is low. In Costa Rica this would make the choice of distribution more critical during the summer months than in the winter, since the frequency of rainfall is higher during the summer season. A close examination of Figures 5-73 to 5-78 shows that the early summer period is one in which the difference between simulated distribution amounts seems to be greatest.

In terms of the individual distributions, the only conclusion that can be drawn at this point is that the Kappa should be removed from further consideration; the primary basis for this is that the Kappa model fails to predict any rainfall at all for the spring months of March and April in several cases. Given the proliferation of other models which adequately match the actual monthly rainfall, there is little reason to further consider the Kappa.

Extreme value analysis. Since the monthly rainfall simulations were inconclusive in terms of distinguishing between the candidate distributions, it was necessary to analyze the behavior of the extreme values that the distributions would yield under repeated circumstances. This provides insight into the nature of the sampling variability for the modeled extreme events.

In general, both the Kappa and Wakeby distributions show sampling variability which is at least as large as distributions with fewer parameters, which indicates that simply choosing larger distributions (i.e., those with more parameters) is not necessarily the best way to enhance the performance of a precipitation model. In fact, the use of the Wakeby or Kappa distributions gave neither the most accurate fit nor the smallest sampling variation. It is likely that this results from an "overfit" of the models, in which too many parameters are being estimated, and demonstrates that parsimony is an important consideration when choosing models.

The Gamma and Pearson III distributions did the best job of fitting the data in all of the cases examined. This lends credence to previous research which has employed the Gamma model for rainfall simulations. In most cases, there is little to distinguish the two distributions in terms of which model to choose. It is intriguing that, despite their mathematical similarities, the two-parameter Gamma often outperforms the three-parameter Pearson III model, indicating that the addition of a location parameter does little more than add an additional layer of complexity to a model that is already adequate for modeling daily rainfall. The only cases when the Pearson III outperforms the Gamma happens in January for Sanatorio Duran and Villa Mills, when relatively few rain events occur. The logical conclusion that can be drawn from this analysis is that the Gamma

distribution should be used in nearly every case, except during the winter in the eastern part of the study region.

Parameter Forecasting

The results from the multiple regression and ARIMA forecast models reiterate the underlying complexity of the rain-generating processes in the region. Even though strong correlations exist between individual atmospheric indices and the parameter estimates, it remains difficult to create models capable of forecasting with any reliability.

The structure and complexity of the multiple regression models, including those with very high multiple correlation coefficients, demonstrates the extreme difficulty in deriving physical relationships from the construction of such models. Many of the models suggest relationships which run counter to previously observed behavior within the system, especially those models which use a large number of NAO measurements as predictors to the exclusion of other variables.

In fact, the extraordinary number of NAO parameters being used in the best-fit models, combined with a marked paucity of NINO34 and SOI measurements, is strongly suggestive of two possible explanations. First, it is possible that the lack of strong individual correlations in the NAO time series allows a greater degree of flexibility in the model-building process. To use an analogy, consider the process of constructing a building: it is easier to construct a wide variety of surfaces using small, non-descript bricks than it is to use very large bricks. The use of smaller shapes allows a greater overall conformance to the finished structure. When constructing multiple regression models, it is rarely a prerequisite that the model predictors must have an absolute physical connection with the dependent variable being modeled; that, of course, would be

the case if perfect relationships existed, but this is certainly not the case in this research.

The fact that the model uses the NAO results more than any other predictor simply reflects the desire for the model to have the maximum amount of variance being explained, regardless of the physical implications. This would also help to explain which the multiple regression model does such a poor job in predicting future parameter values: the lack of physical connectivity between the predictors and the predictand results in highly unstable predictions.

The second possible explanation is that there is a considerable amount of underlying physical connection between the NAO and the other indices which does not directly result in an immediately recognizable correlation pattern. It is possible that such patterns would only be visible when the other variables are being considered simultaneously, which occurs during the model-building process of multiple regression.

It is clear, however, that whichever explanation is best suited in the present case, the ARIMA model does a better job of predicting the future parameter values than the multiple regression model. This is probably due to the highly seasonal nature of the parameters; an ARIMA model can easily take seasonality into account when formulating the necessary autoregressive and moving average parameters. It is important to note, as well, that the multiple regression model required the creation of individual models for each month, and in several cases the model was quite large (over 10 values being used to constitute the model), whereas the ARIMA model required the estimation of only two parameters: a seasonal autoregressive parameter and a seasonal moving average parameter. The results indicate that the ARIMA approach is better suited to forecasting parameter values, both from the standpoint of total accuracy and parsimony.

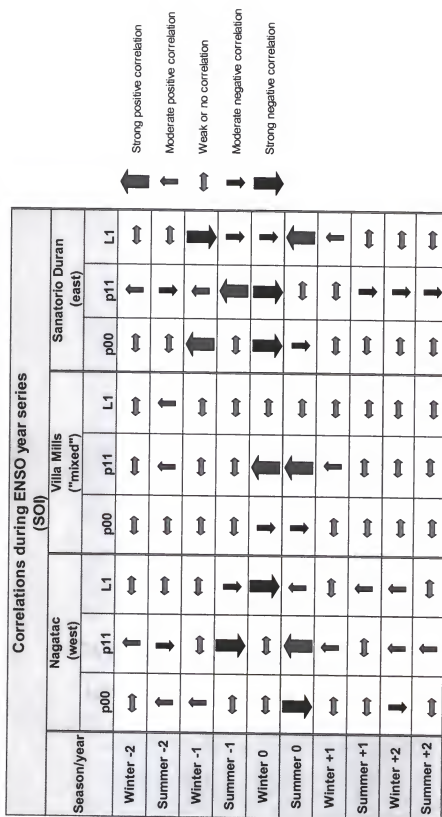


Figure 6-1. Summary of SOI correlations.

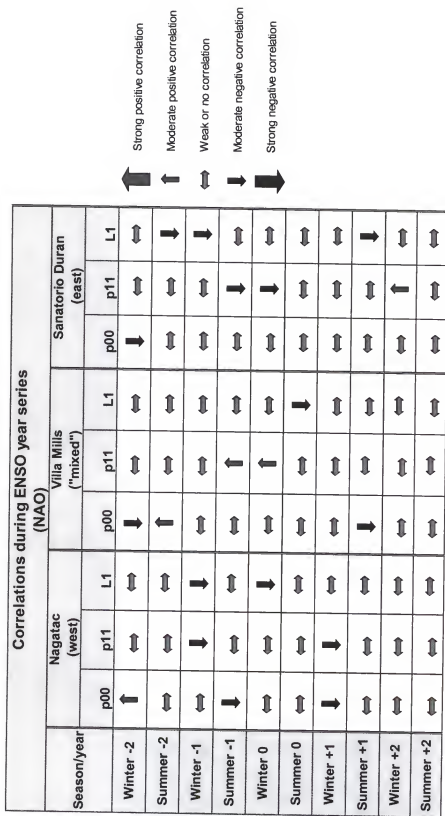


Figure 6-2. Summary of NAO correlations.

Correlations during ENSO year series (NINO34)												
Season/year	Nagatac (west)			Villa Mills ("mixed")			Sanatorio Duran (east)					
	p00	p11	L1	p00	p11	L1	p00	p11	L1			
Winter -2	↔	↓	↓	↔	↔	↓	↔	↔	↔	↔	↔	↔
Summer -2	↓	↓	↑	↔	↓	↔	↓	↔	↓	↓	↓	↓
Winter -1	↔	↔	↓	↓	↑	↔	↓	↔	↔	↔	↔	↔
Summer -1	↓	↓	↑	↓	↔	↔	↔	↓	↔	↓	↓	↓
Winter 0	↔	↔	↓	↑	↓	↔	↓	↓	↓	↓	↓	↓
Summer 0	↓	↓	↓	↔	↓	↓	↓	↓	↓	↓	↓	↓
Winter +1	↑	↓	↑	↔	↓	↔	↑	↔	↔	↔	↔	↔
Summer +1	↓	↔	↔	↔	↓	↔	↓	↑	↔	↔	↔	↔
Winter +2	↔	↔	↓	↔	↑	↔	↔	↔	↔	↔	↔	↔
Summer +2	↓	↑	↔	↑	↔	↔	↔	↑	↔	↔	↔	↔

 Strong positive correlation
 Moderate positive correlation
 Weak or no correlation
 Moderate negative correlation
 Strong negative correlation

Figure 6-3. Summary of NINO34 correlations.

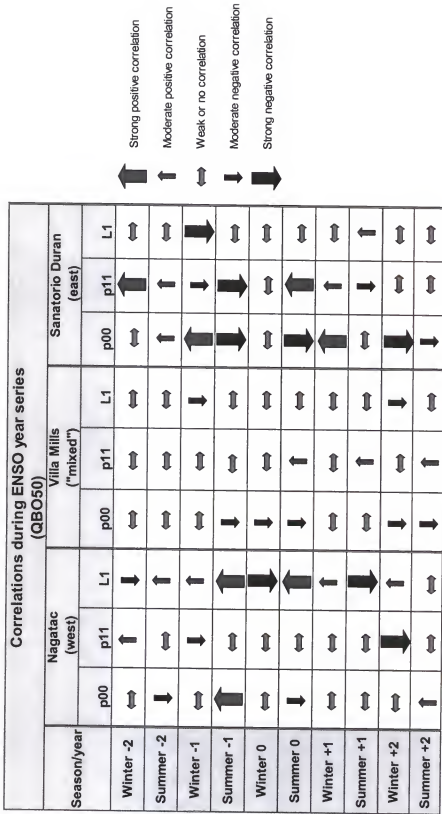


Figure 6-4. Summary of QBO50 correlations.

CHAPTER 7 CONCLUSION

One of the driving ideas behind this research has been the desire to bridge the gap between the understanding of large-scale climate dynamics and the immediate practical small-scales of the end-user. The past few years have seen a rapid increase in the ability of researchers to create reasonable forecasts of processes at global and hemispheric scales, especially as it relates to the evolution of ENSO-related climatic events. At the same time, the need for information at the regional and local levels of agencies and planners has seen a dramatic increase as governments attempt to better utilize their available resources. The results of this study help to provide much-needed insight into the application of large-scale climate prediction to end-users, who can immediately apply these findings to problems at the local level.

While this research has focused exclusively on the characterization of Costa Rican daily rainfall, the results provide important tools for understanding similar problems in many other areas of the globe. It is within that context that it points towards some important implications and conclusions that have broad applicability, both to the general field of statistical climatology and particularly to the specific case of the modification of daily tropical precipitation by the El Niño-Southern Oscillation phenomenon. The conclusions can be grouped into three broad categories of results: those dealing with specific statistical modeling issues, those dealing with climatological

issues, and those dealing with specific applications of the results to other human endeavors.

Statistical Modeling

The examination of precipitation characteristics in much of the equatorial/tropical world is hampered by a comparative lack of available data, which is much more readily available in the more developed temperate areas of the globe. For this reason, it has been unclear whether the techniques developed for more climatologically temperate (and data-rich) countries would be adequate to describe the statistical nature of daily rainfall. The present research demonstrates that techniques exist that can help to mitigate the effects of sparse data, a problem which plagues many of the world's tropical regions.

Statistical Modeling — Occurrence

Indications are that the preferred order for a Markov model describing the occurrence of daily rainfall cannot be assigned *a priori* based on findings from similar research in temperate regions. The choice of model order is an important consideration, and reflects the degree to which the occurrence process can be considered a serially-dependent (i.e., autocorrelated) process. Previous research has concluded that a first order model is the most efficient model for nearly all locations; the present research suggests greater complexity in which the selected order varies with season and rainfall-generating process. Further complicating matters, Katz and Parlange (1998) point towards a different conclusion for temperate data in California, the results of which may be pertinent in the present case; they state that the lower-order (usually first-order) models currently used may contribute to "overdispersion" of simulated data, i.e., an increase in rainfall variance, probably due to the loss of information when low-order models are used.

This conclusion implies that the choice of model order for future research and applications must be thoroughly tested for each location. The choice will generally fall on the ultimate use of the model, and the degree to which the end-user can tolerate the variance increase of a low-order model versus the decrease in parsimony associated with higher-order models. In either case, this research points towards the need for more research being performed in a wider variety of geographic regions to clarify the impacts of order choice on the modeling results.

Statistical Modeling — Intensity

The formulation of an appropriate statistical intensity model has been a contentious issue, both in the way in which the best-fit distribution are selected, and also in the manner in which the distributional parameters are calculated. Both topics have been addressed, and methods have been presented that are immediately applicable to a variety of situations for which long-term data may not be available, or when the data are heavily skewed by very large daily rainfalls, conditions which prevail in many humid tropical regions.

Utilization of L-moments. The dual problems of distribution choice and parameter estimation are intertwined; testing of distributional adequacy is dependent upon parameter estimation, which itself requires assumptions about the distributional form. The use of L-moments greatly simplifies this process. The moments are reflective of the underlying properties of the daily rainfall intensity characteristics, and produce distribution parameters which are much more robust in the face of extreme rainfall values than traditional statistical moments. This is particularly true in cases where the data are limited; they provide a much higher level of confidence in the higher-order moments

which are calculated in these situations. They are also very easy to calculate and interpret, and quickly yield robust parameter estimates for a variety of distributions.

Distribution choice. Conclusions regarding the choice of the distribution used to model daily rainfall intensities appear somewhat counter-intuitive. Contrary to expectations, distributions with larger numbers of parameters, which are thought to allow a greater degree of adjustment across the entire range of values, do not necessarily give the best fit. While some degree of adjustment in the skew of the distribution is necessary (thus eliminating such distributions as the Normal and Exponential from consideration), simply adding more parameters does not gain much advantage when fitting the highly right-skewed data associated with daily rainfall totals. The danger of “overfitting” the model is clearly illustrated in this research, especially when the underlying assumptions that govern the behavior of the model are unclear.

In nearly every case examined, the distribution of choice was either the two-parameter Gamma or the three-parameter Pearson III model. Both distributions evolve from the same general function (a four-parameter Generalized Gamma or Hypergamma distribution (Evans et al., 1993, page 80)), could serve as a useful “parent” distribution for both the Gamma and Pearson III. To date, however, the L-moment characteristics of the Hypergamma distribution have not been explored, so this remains a fruitful topic of future study.

Climatology

General Costa Rican Climatology

Conclusions concerning the complex spatial and temporal nature of Costa Rica’s precipitation climatology, previously only detailed at coarser time scales, appear to be

supported by this research, and provide a physically-based, and meaningful, interpretation of the statistical models. Distinctive signatures include both the *veranillos* and *nortes* in the p_{xx} and L_1 graphs, which further elucidate the nature of the mechanisms which give rise to these seasonal changes. These parameters indicate that, during the mid-summer *veranillos* especially, the Trade winds exert a greater influence over the magnitude of daily rainfall than over the frequency of rainfall events in the eastern part of Costa Rica. Changes in the Trades are more likely to be reflected in how much rain falls during a storm rather than how often the storms occur, whereas such changes are brought about by changes in both the size and frequency of events along the Pacific flank. This suggests that the Trades represent an ever-present rainfall-generating mechanism which varies in intensity throughout the year. The ITCZ, by contrast, is present only during the boreal summer, and its northernmost migration changes according to general atmospheric conditions; it acts more in an "on/off" fashion for the western part of Costa Rica, its presence being the driving mechanism for both rainfall occurrence and intensity. The seasonal effects of the ITCZ extend beyond the topographic limits of the Pacific slope, representing a powerful rain-inducing mechanism, whose impact extends well into the Reventazón basin of the Central Valley and headwaters of other rivers which flow towards the Caribbean bringing Pacific-like conditions during the late summer. The *nortes*, by contrast, are limited in extent to the eastern regions of the country, where changes in the intensity of rain events predominates over changes in the occurrence.

Relationships with ENSO

Costa Rica constitutes an excellent location in which to study of the effects of ENSO in the tropics, lying between both the Atlantic and Pacific Oceans, and connecting

the extra-tropical and tropical American continents. This is of particular importance when the influence of extra-tropical processes on tropical rainfall are of interest. Different regions of the country can experience flooding and droughts at various times of year as a result of ENSO. This research provides increased insight into the pivotal role of ENSO in the interannual variation of daily rainfall characteristics in various regions of Costa Rica.

Changes in daily precipitation characteristics brought about by ENSO are consistent with the “canonical” El Niño, in that the alterations the daily precipitation climatology can be directly traced to the pressure changes in the equatorial Pacific and the associated advection of warm waters towards the eastern boundary of the Pacific basin, which herald the increase in the Trade winds across the Central American isthmus. The daily rainfall record clearly records the increase in precipitation intensity in the eastern part of the study area, and the western areas experience a decrease in rainfall in the same time period. This illustrates the connection between the characteristics of the small-scale rainfall changes and the larger climatological issue of ENSO evolution.

Segregated year analysis. The El Niño-Southern Oscillation clearly has a significant impact on the year-to-year variation of the climatology of Costa Rica. This impact is most noticeable during the summer months, when the Trades and the ITCZ are at their maximum intensities. The occurrence of an El Niño event is clearly reflected during the early summer as a general increase in the intensity of rains brought by the Trades in the eastern part of the study area, indicating an increase in the intensity of the Trades themselves.

The interaction between the Trades and the ITCZ during mid-summer, when the Trades increase in intensity, restrain the northward migration of the ITCZ and give rise to

the *veranillos* along the Pacific slope, shows a clear relationship with the state of the ENSO. During El Niño phases, the *veranillos* are clearly extended in duration and intensity; the probabilities of extended wet periods are significantly reduced, as is the quantity of rainfall delivered in any rainy day. El Niño causes both the size and duration of storms in the west to be reduced during this period in western Costa Rica.

La Niña conditions are at least as important as the El Niño in determining the character of daily rainfall across Costa Rica. La Niña years give rise to considerably higher rainfall intensities during the late summer, which is reflective of the increase in the strength and northward positioning of the ITCZ. This is accompanied by a significant decrease in the length of the dry spells, indicating that the ITCZ is promoting the development of more storms on a daily basis. By contrast, the Caribbean flank experiences lower rainfall intensities during the same period, demonstrating that the Trade winds have weakened compared to El Niño years.

The ENSO is also a major influence in the development of the *nortes*. The early and late season *nortes* react differently to the state of the ENSO. The early season *nortes* (December/January) are more likely to experience a change in the probability of rainfall occurrence, without a similar change in rainfall amounts. El Niño years have significantly longer dry spells and shorter wet spells, while La Niña years see a reversal in this trend; this results in El Niño years having drier than normal conditions during early winter, and La Niña years having wetter conditions, both due to changes in how often it rains. By contrast, the late winter (March) *nortes* are primarily modified in terms of their rainfall intensities, rather than occurrence, and this is most apparent during La Niña years. Thus, El Niño year winters are drier overall because of early season changes in rainfall

occurrence, while La Niña years are moister because of both occurrence and intensity changes.

Lag-correlation analysis. Three important conclusions arise from the lag-correlation analysis. First, it appears that the conditions in the Atlantic basin, as measured by the NAO, are of minimal importance in the characterization of daily rainfalls. The lack of any significant pattern of strong correlations between the NAO and the daily rainfall parameters implies that the perceived “pushing” effect of the North Atlantic high pressure area is negligible for Costa Rica. This contradicts what has been seen in terms of the Trade wind/rainfall connections; a possible explanation of this is the problem of using what is essentially a measure of extreme North Atlantic climate to represent Trade wind strength. This further implies that the important connection is between Costa Rica and the equatorial Pacific, with little influence being placed on the rainfall by teleconnections between the equatorial Pacific and other geographic areas of the Earth’s ocean-atmosphere system.

Second, the role of the quasi-biennial oscillation is an important factor, more so in the extreme eastern and western parts of the study area rather than in the transitional region towards the center. The reasons for this association are as yet not well understood, but, as noted earlier, similar relationships have been reported in other areas of the world. The state of the QBO clearly represents an indication of significant energy transfer between the troposphere and stratosphere, but unfortunately the exact mechanisms are unclear.

Third, the results indicate that the El Niño phenomenon represents, to a greater degree than is usually thought, a continuum of processes, rather than a simple event.

These are felt well before, and after, the sea surface temperatures reach their maximum off-shore of South America. This implies a consideration of the ENSO cycle as an evolutionary process, rather than as a singular event.

In addition, it seems clear that the QBO cycle acts on a different time scale than the ENSO, yet they both have a significant association with the daily rainfall characteristics. This leads to the conclusion that the ENSO and the QBO are probably experiencing a degree of interaction with one-another, which in turn is likely to have effects on Central American rainfall that differs in time scale from either the “pure” ENSO or QBO process. It is likely that the superimposed periodicities of geophysical processes such as the QBO and ENSO would have the effect of modulating the strength of an El Niño event at time scales quite different from either of the underlying events.

Applications to Other Areas

One of the most important results to come out of this research is the idea that knowledge of the ENSO process, and its effects on the daily climatology of the region, can enhance the predictability of rainfall at larger time scales, both in terms of the amount of rainfall expected and its daily characteristics. The ability to forecast the statistical parameters that describe daily rainfall using only a small set of atmospheric indicators can be of immeasurable importance to farmers, engineers, planners, and government agencies.

Simulations and forecasts of daily rainfall characteristics based upon the state of the ENSO is an entirely reasonable expectation, and well within the grasp of currently utilized statistical practices. Simulated rainfalls show a very high correspondence with the historical averages, based upon the state of the El Niño. This helps to illustrate the

dependence of Costa Rican rainfall on the El Niño process. Forecasts tend to be rather less accurate, which further demonstrates the complexity of the climatological system; clearly processes other than just the ENSO are coming into play. In addition, the lack of complete understanding of the underlying physical connections makes it difficult to construct physically meaningful forecast models.

Despite the complexity of the Costa Rican climatological system, and the inherent unpredictability of the El Niño-Southern Oscillation phenomenon, it is clear that the system is comprehensible at a variety of temporal and physical scales that are important to researchers, planners, and agencies. The ability to mathematically model the effects of a global process such as ENSO on the climatology of Costa Rica presents the opportunity to create models and forecasts that can be utilized at numerous levels. This research helps to bridge the gap between the high-level research of atmospheric scientists and the individuals who must make pragmatic and immediate decisions at the local level, as well as providing a framework for the further study of tropical precipitation climatology.

APPENDIX L-MOMENTS

The use of L-moments in hydroclimatology has been established by Hosking (1989, 1990), Hosking and Wallis (1987, 1993), Guttman (1993), and Vogel and Fennessey (1993) as a superior means of understanding the underlying distributional characteristics of a time series of data when the underlying distributional form is unknown *a priori*. The primary utility of L-moments is the stability of the resulting moments in the face of highly skewed data, such as daily rainfall data from a humid tropical environment as is found in Costa Rica.

A demonstration using the data from the three stations from Sanatorio Duran, Villa Mills, and Nagatac is presented in this section. The formulas presented in the Methodology section are employed to determine the standard product moments and corresponding L-moments for the aggregate of each month of data from each stations. In this way, 36 station-months of data were separately analyzed. The resulting product moment and L-moment coefficients of variation and coefficients of skew were plotted for the full data sets, and for the same data sets with the single largest daily rainfall removed. The changes in the moment values were also plotted.

When the standard product moments for the full data sets and the abbreviated data sets are compared, it is clear that substantial changes in the moments have resulted due to the loss of a single largest value. This translates to a large potential change in the

evaluation of the underlying distribution; put another way, the use of standard product moments has resulted in a substantial degree of instability in the distribution which describes the time series. The L-moment results show much less change in the resulting moments, which in turn gives a much more robust estimate of the underlying distribution.

Product Moments

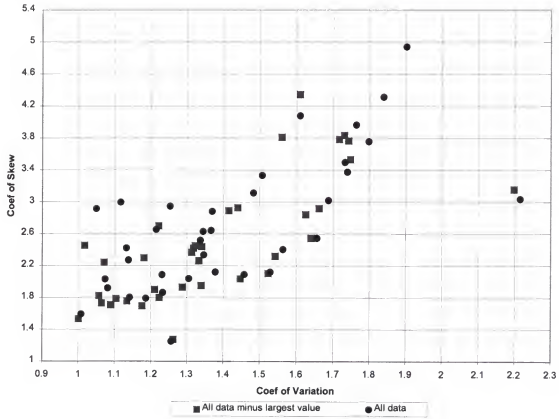


Figure A-1. Product moment results

Linear Moments

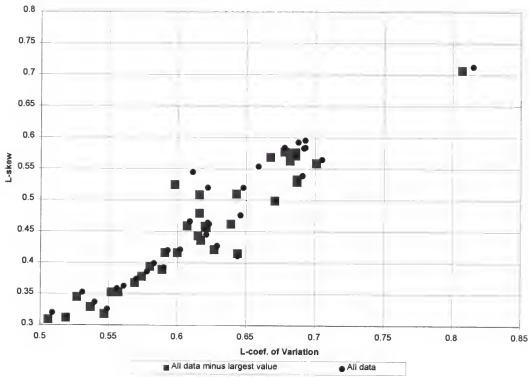
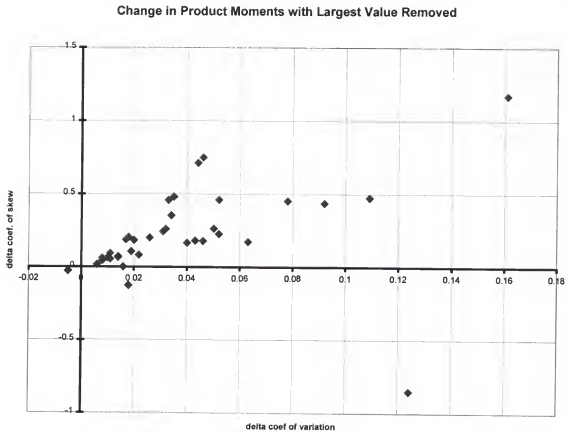


Figure A-2. L-moment results



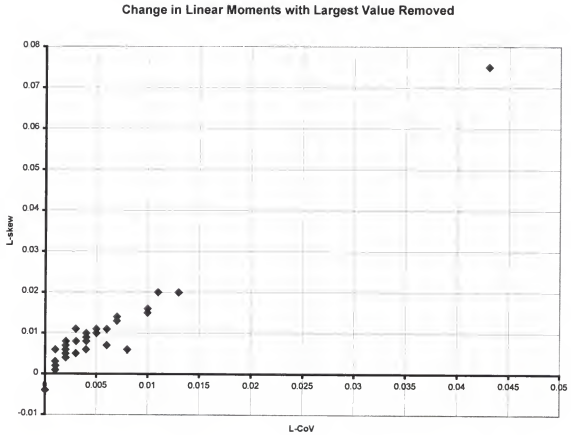


Figure A-4. Change in L-moments with largest value removed.

REFERENCES

Akaike, H., 1974. "A new look at the statistical model identification," IEEE Transactions on Automatic Control, AC-19, 716-723.

Aldenderfer, M.S. and R.K. Blashfield, 1984. Cluster Analysis. SAGE University Paper Series on Quantitative Applications in the Social Sciences, series 07-044. Beverly Hills, CA: SAGE Publications, 87 pp.

Baumgartner, T.R. and N. Christensen, 1985. "Coupling of the Gulf of California to large scale interannual climate variability," Journal of Marine Research, 43, 825-848.

Berlage, H.P., 1957. "Fluctuations of the general atmospheric circulation of more than one year, their nature and prognostic value," K. Ned. Meteorol. Inst. Meded. Verh., 69, 152.

Bjerknes, J., 1966. "A possible response to the atmospheric Hadley circulation to equatorial anomalies of ocean temperature," Tellus, 18, 820-829.

Bjerknes, J., 1969. "Atmospheric teleconnections from the equatorial Pacific," Monthly Weather Review, 97, 163-172.

Bryson, R.A. and P.M. Khun, 1961. "Stress-differential induced divergence with application to littoral precipitation." Erkunde, 15, 287-294.

Cane, M.A. and P.R. Gent, 1977. "Reflection of low-frequency equatorial waves at arbitrary western boundaries," Journal of Marine Research, 42, 395-432.

Caskey, J.E., 1963. "A Markov chain model for the probability of precipitation occurrence in intervals of various lengths," Monthly Weather Review, 91, 298-301.

Cavazos, T. and S. Hastenrath, 1990. "Convection and rainfall over Mexico and their modulation by the Southern Oscillation," International Journal of Climatology, 10, 377-386.

Caviedes, C.N., 1973. "Secas and El Niño: Two simultaneous climatical hazards in South America," Proceedings of the Association of American Geographers, 5, 44-49.

Caviedes, C.N., 1975. "El Niño 1972: Its climatic, ecological, human and economic implications," The Geographical Review, 65, 439-509.

Caviedes, C.N., 1981. "Rainfall in South America: Seasonal trends and spatial correlations," Erkunde, 35, 107-118.

Caviedes, C.N., 1984. "El Niño 1982-83," The Geographical Review, 74, 267-290.

Chacon, R.E. and W. Fernandez, 1985. "Temporal and spatial rainfall variability in the mountainous region of the Reventazon River basin, Costa Rica," Journal of Climatology, 5, 175-188.

Chatterjee, S. and B. Price, 1991. Regression Analysis By Example. New York, NY: John Wiley and Sons, Inc.

Chin, E.H., 1977. "Modeling daily precipitation occurrence process with Markov chain," Water Resources Research, 13, 949-956.

Chu, P.-S. and R.W. Katz, 1985. "Modeling and forecasting the Southern Oscillation: A time-domain approach," Monthly Weather Review, 113, 1876-1888.

Chu, P.-S. and R.W. Katz, 1989. "Spectral estimation from time series models with relevance to the Southern Oscillation," Journal of Climate, 2, 86-90.

Coe, R. and R.D. Stern, 1982. "Fitting models to daily rainfall data," Journal of Applied Meteorology, 21, 1024-1031.

Cooley, J.W. and J.W. Tukey, 1965. "An algorithm for the machine calculation of complex Fourier series," Mathematical Computations, 19, 297-301.

Cong, S. and Y. Li, 1993. "Identification of the underlying distribution form of precipitation by using regional data," Water Resources Research, 29, 1103-1111.

Craven, P. and G. Wahba, 1979. "Smoothing noise data with spline functions," Numerical Mathematics, 31, 377-403.

D'Agostino, R.B. and M.A. Stephens, 1986. Goodness-of-Fit Techniques. New York, New York: Marcel Dekker, Inc.

De Boor, C., 1978. A Practical Guide to Splines. New York, N.Y.: Springer-Verlag.

Deser, C. and J.M. Wallace, 1987. "El Niño events and their relation to the Southern Oscillation," Journal of Geophysical Research, 92, 14189-14196.

Diaz, H.F. and G.N. Kiladis, 1992. "Atmospheric teleconnections associated with the extreme phase of the Southern Oscillation," in El Niño: Historical and Paleoclimatic

Aspects of the Southern Oscillation (H.F. Diaz and V. Markgraf, eds.). Cambridge, England: Cambridge University Press.

Diaz, H.F. and R.S. Pulwarty, 1994. "An analysis of the time scales of variability in centuries-long ENSO-sensitive records in the last 1000 years," Climatic Change, 26, 317-342.

Dutton, J.A., 1986. The Ceaseless Wind: An Introduction to the Theory of Atmospheric Motion. New York: Dover Publications, Inc., 617 pp.

Elderton, W.P., 1953. Frequency Curves and Correlations, Fourth Edition. Washington, D.C.: Harren Press.

Emory, W.J. and K. Hamilton, 1985. "Atmospheric forcing of interannual variability in the northeast Pacific Ocean: Connections with El Niño," Journal of Geophysical Research, 90, 857-868.

Enfield, D.B., 1989. "El Niño, past and present," Review of Geophysics, 27, 159-187.

Enfield, D.B., 1992. "Historical and prehistorical overview of El Niño/Southern Oscillation," in El Niño: Historical and Paleoclimatic Aspects of the Southern Oscillation (H.F. Diaz and V. Markgraf, eds.). Cambridge, England: Cambridge University Press.

Epstein, E.S., 1992. "ENSO's precipitation signature in the U.S.: Are wet days wetter or more frequent?" Proceedings of the Sixteenth Annual Climate Diagnostic Workshop, held at the Department of Atmospheric Sciences, University of California, Los Angeles, CA, March 1992.

Estoque, M.A., J. Luque, M. Chandeck-Monteza and J. Garcia, 1985. "Effects of El Niño on Panama rainfall," Geofisica Internacional, 24, 355-381.

Evans, M., N. Hastings, and B. Peacock, 1993. Statistical Distributions. New York, NY: John Wiley and Sons, Inc.

Fernández, W. and P. Ramírez, 1991. "La Osilacion del Sur sus efectos en Costa Rica: Una revision," Tecnologia en Marcha, 11, 3-11.

Friedman, J.A., 1991. "Multivariate adaptive regression splines," Annals of Statistics, 19(1), 1-141.

Friedman, J.A. and B.W. Silverman, 1989. "Flexible parsimonious smoothing and additive modeling," Technometrics, 31(1), 3-39.

George, R.K., P. Waylen, and S. LaPorte, 1998. "Interannual variability of annual streamflow and the Southern Oscillation in Costa Rica." Hydrological Sciences Journal, 43(3), 1-18.

Gill, A.E. and E.M. Rasmussen, 1983. "The 1982-1983 climate anomaly in the equatorial Pacific," Nature, 306, 229-234.

Gleick, P.H., 1993. Water in Crisis: A Guide to the World's Fresh Water Resources. New York, New York: Oxford University Press, 473 pp.

Goldberg, R.A., M. Tisnado and R.A. Scofield, 1987. "Characteristics of extreme rainfall events in northwestern Peru during the 1982-1983 period," Journal of Geophysical Research, 92, 14225-14241.

Goode's World Atlas, 18th Edition, 1992, (Edward B. Espenshade, Jr., editor). New York, New York: Rand McNally, 368 pp.

Gray, W.M., 1984a. "Atlantic seasonal hurricane frequency. Part I: El Niño and the 30 mb quasi-biennial oscillation influences," Monthly Weather Review, 112, 1649-1668.

Gray, W.M., 1984b. "Atlantic seasonal hurricane frequency. Part II: Forecasting its variability," Monthly Weather Review, 112, 1669-1683.

Greenwood, J.A., J.M. Landwehr, N.C. Matalas and J.R. Wallis, 1979. "Probability weighted moments: definitions and relation to parameters of several distributions expressible in inverse form," Water Resources Research, 15, 1049-1054.

Gumbel, E.J., 1958. Statistics of Extreme Values. New York, NY: Columbia University Press.

Guttman, N.B., 1993. "The use of L-moments in the determination of regional precipitation climates," Journal of Climate, 6, 2309-2325.

Guttman, N.B., 1994. "On the sensitivity of sample L-moments to sample size," Journal of Climate, 7, 1026-1029.

Guttman, N.B.; J.R.M. Hosking and J.R. Wallis, 1993. "Regional precipitation quantile values for the continental United States computed from L-moments," Journal of Climate, 6, 2326-2340.

Guzman, A.G. and W.C. Torres, 1985. "Daily rainfall probabilities: Conditional upon prior occurrence and amount of rain," Journal of Climate and Applied Meteorology, 24, 1009-1014.

Haan, C.T., 1977. Statistical Methods in Hydrology. Ames, Iowa: The Iowa State University Press.

Halpert, M.S. and C.F. Ropelewski, 1992. "Surface temperature patterns associated with the Southern Oscillation," Journal of Climate, 5, 577-593.

Hamilton, K., 1990. "A look at the recently proposed solar-QBO-weather relationship," Journal of Climate, 3, 497-503.

Hamilton, K. and R.R. Garcia, 1986. "El Niño/Southern Oscillation events and their associated mid-latitude teleconnections," Bulletin of the American Meteorological Society, 67, 1354-1361.

Hansen, D.V. and C.A. Paul, 1984. "Genesis and effects of long waves in the equatorial Pacific," Journal of Geophysical Research, 89, 10431-10440.

Harrison, D.E. and P.S. Schopf, 1984. "Kelvin-wave-induced anomalous advection and the onset of surface warming in El Niño events," Journal of Physical Oceanography, 14, 923-933.

Harrison, J.M., 1995. "The use of L-moments in the identification of probability distributions in climatological data." Presented at the 91st Annual Meeting of the Association of American Geographers, March 14-18, 1995, Chicago, Illinois.

Harrison, J.M. and P.R. Waylen, 1996. "The regionalization of precipitation patterns in Central America using cluster analysis." Presented at the 92nd Annual Meeting of the Association of American Geographers, April 9-13, 1996, Charlotte, North Carolina.

Hastenrath, S.L., 1976. "Variations in low latitude circulation and extreme climatic events in the tropical Americas," Journal of the Atmospheric Sciences, 33, 202-215.

Hastenrath, S.L., 1990. "Diagnostics and prediction of anomalous river discharge in northern South America." Journal of Climate, 3, 1080-1096.

Hirschboeck, K.K., 1988. "Flood hydroclimatology," in Flood Geomorphology (V.C. Baker, R.C. Kochel and P.C. Patton, eds). New York: Wiley-Interscience, 503 pp.

Horel, J.D. and J.M. Wallace, 1981. "Planetary-scale atmospheric phenomena associated with the Southern Oscillation," Monthly Weather Review, 109, 813-829.

Horn, L.H. and R.A. Bryson, 1960. "Harmonic analysis of the annual march of precipitation over the United States," Annals of the Association of American Geographers, 50(2), 157-171.

Hosking, J.R.M., 1985. "Maximum-likelihood estimation of the parameters of the generalized extreme value distribution," Applied Statistics, Statistical Algorithm AS 215, 301-310.

Hosking, J.R.M., 1989. The Theory of Probability Weighted Moments. IBM Technical Report RC-12210. Yorktown Heights, New York: IBM Corporation, 160 pp.

Hosking, J.R.M., 1990. "L-moments: Analysis and estimation of distributions using linear combinations of order statistics," Journal of the Royal Statistical Society B, 52, 105-124.

Hosking, J.R.M. and J.R. Wallis, 1987. "Parameter and quantile estimation for the generalized Pareto distribution," Technometrics, 29, 339-349.

Hosking, J.R.M. and J.R. Wallis, 1993. "Some statistics useful in regional frequency analysis," Water Resources Research, 29, 271-281.

Hughes, J.P. and Guttorp, P., 1994. "A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena," Water Resources Research, 30, 1535-1546.

Jenkinson, A.F., 1954. "The frequency distribution of the annual maximum (or minimum) values of meteorological elements," Quarterly Journal of the Royal Meteorological Society, 87, 158-171.

Jones, P.G. and P.K. Thornton, 1993. "A rainfall generator for agricultural applications in the tropics," Agricultural and Forest Meteorology, 63, 1-19.

Julian, P.R. and R.M. Chervin, 1978. "A study of the Southern Oscillation and Walker Circulation phenomenon," Monthly Weather Review, 106, 1433-1451.

Katz, R.W., 1977. "Precipitation as a chain-dependent process," Journal of Applied Meteorology, 16, 671-676.

Katz, R.W., 1981. "On some criteria for estimating the order of a Markov chain," Technometrics, 23, 243-249.

Katz, R.W. and M.B. Parlange, 1993. "Effects of an index of atmospheric circulation on stochastic properties of precipitation," Water Resources Research, 29, 2335-2344.

Katz, R.W. and M.B. Parlange, 1998. "Overdispersion phenomenon in stochastic modeling of precipitation," Journal of Climate, 11, 591-601.

Kazadi, S.N. and F. Kaoru, 1996. "Interannual and long-term climate variability over the Zaire River basin during the last 30 years," Journal of Geophysical Research-Atmospheres, 101(D6), 21351-21360.

Keen, R.A., 1982. "The role of cross-equatorial cyclone pairs in the Southern Oscillation," Monthly Weather Review, 110, 1406-1416.

Keppenne, C.L. and M. Ghil, 1992. "Adaptive filtering and prediction of the Southern Oscillation Index," Journal of Geophysical Research, 97(D18), 20449-20454.

Kirtman, B.P. and S.E. Zebiak, 1997. "ENSO simulation and prediction with a hybrid coupled model," Monthly Weather Review, 125, 2620-2641.

Knisel, W.G., K.G. Renard, and L.J. Lane, 1979. "Hydrologic data collection: How long is long enough?" Irrigation and Drainage in the Nineteen-Eighties, ASCE Special Conference Transactions, Albuquerque, NM, July 17-20, 1979.

Knox, R.A. and D. Halpern, 1982. "Long range Kelvin wave propagation of transport variation in the Pacific Ocean equatorial currents," Journal of Marine Research, 40 (sup), 329-339.

Kodera, K., 1991. "The solar and equatorial QBO influences on the stratospheric circulation during the early northern-hemisphere winter," Geophysical Research Letters, 18(6), 1023-1026.

Labitzke, K. and H. van Loon, 1993. "Some recent studies of probable connections between solar and atmospheric variability," Annales Geophysicae, 11, 1084-1094.

Lall, U. and M. Mann, 1995. "The Great Salt Lake: A barometer of low-frequency climatic variability," Water Resources Research, 31(10), 2503-2515.

Lall, U., B. Rajagopalan, and D.G. Tarbotan, 1996. "A nonparametric wet/dry model for resampling daily precipitation," Water Resources Research, 32, 2803-2823.

Lall, U., T. Sangoyomi and H.D.I. Abarbanel, 1996. "Nonlinear dynamics of the Great Salt Lake: Nonparametric short-term forecasting," Water Resources Research, 32(4), 975-985.

Lall, U., D.P. Ames, and C. Baldwin, 1998. ItsMARS version 1.6. Utah State University, Logan, UT. WWW address: <http://pub.uwrl.usu.edu/~ulall/code>.

Landwehr, J.M., N.C. Matalas and J.R. Wallis, 1979. "Estimation of parameters and quantiles of Wakeby distributions: 1. Known lower bounds," Water Resources Research, 15, 1361-1379.

Mann, M.E., U. Lall and B. Saltzman, 1995. "Decadal-to-centennial-scale climate variability: Insights into the rise and fall of the Great Salt Lake," Geophysical Research Letters, 22(8), 937-940.

Mann, M.E. and J. Park, 1993. "Spatial correlations of interdecadal variation in global surface temperature," Geophysical Research Letters, 20(11), 1055-1058.

Mann, M.E. and J. Park, 1994. "Global-scale modes of surface temperature variability on interannual to century timescales," Journal of Geophysical Research, 99(D12), 25819-25833.

Markovic, R.D., 1965. "Probability functions of best fit to distributions of annual precipitation and runoff," Hydrology Paper 8, Ft. Collins, Colorado: Colorado State University.

Mason, S.J. and P.D. Tyson, 1992. "The modulation of sea surface temperatures and rainfall associations over southern Africa with solar activity and the Quasi-Biennial Oscillation," Journal of Geophysical Research, 97(D5), 5847-5856.

Mason, S.J., P.R. Waylen, G.M. Mimmack, B. Rajaratnam and J.M. Harrison, 1996a. "Estimation of rainfall return periods over South Africa," Water South Africa, in review.

Mason, S.J.; P.R. Waylen, G.M. Mimmack, B. Rajaratnam and J.M. Harrison, 1996b. "Changes in flood frequencies in South Africa," International Journal of Climatology, in review.

Mitra, K., S. Mukherji, and S.N. Dutta, 1991. "Some indications of 18.6 year luni-solar and 10-11 year solar cycles in rainfall in north-west India, the Plains of Uttar Pradesh and north-central India," International Journal of Climatology, 11, 645-652.

Mooley, D.A. and H.L. Crutcher, 1968. An application of the gamma distribution function to Indian rainfall. Environmental Sciences Services Administration Technical Report EDS-5, U.S. Department of Commerce, Silver Springs, Maryland.

NOAA, 1996. "CPC -- Data: Current Monthly Atmospheric and SST Index Values." NOAA Climate Prediction Center World-Wide Web Site: <http://nic.fb4.noaa.gov/data/cddb>.

NOAA, 1997. "Comprehensive Ocean-Atmosphere Data Set." NOAA Climate Diagnostic Center World-Wide Webb Site: <http://www.cdc.noaa.gov/coads>.

Norman, G.R. and D.L. Streiner, N., 1986. PDQ Statistics. Toronto, Canada: B.C. Decker Inc.

Ochoa, N. and O. Gomez, 1987. "Dinoflagellates as indicators of water masses during El Niño, 1982-1983," Journal of Geophysical Research, 92, 14347-14354.

Periago, M.C., X. Lana, C. Serra and G.F. Mills, 1991. "Precipitation regionalization: An application using a meteorological network in Catalonia (NE Spain)," International Journal of Climatology, 11, 529-543.

Perry-Castanada Map Collection, 1987. Costa Rica. Austin, TX: The University of Texas at Austin. http://www.lib.utexas.edu/Libs/PCL/Map_collection/Map_collection.html.

Philander, S.G., 1981. "The response of equatorial oceans to a relaxation of the trade winds," Journal of Physical Oceanography, 11, 176-189.

- Philander, S.G., 1989. El Niño, La Niña, and the Southern Oscillation. New York, New York: Academic Press, 291 pp.
- Pickering, N.B., J.W. Hansen, J.W. Jones, C.M. Wells, V.K. Chan and D.C. Godwin, 1994. "Weatherman: A utility for managing and generating daily weather data," Agronomy Journal, 86(2), 332-337.
- Poveda, G. and O. Mesa, 1997. "Feedbacks between hydrological processes in tropical South America and large scale oceanic-atmospheric phenomena," Journal of Climate, 10, 2690-2702.
- Preisendorfer, R.W., 1988. Principle Component Analysis in Meteorology and Oceanography. New York, New York: Elsevier Science Publishers, 425 pp.
- Rasmussen, E.M. and T.H. Carpenter, 1983. "The relationship between eastern equatorial Pacific sea surface temperatures and rainfall over India and Sri Lanka," Monthly Weather Review, 111, 517-28.
- Ropelewski, C.F. and M.S. Halpert, 1987. "Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation," Monthly Weather Review, 115, 1606-26.
- Ropelewski, C.F., M.S. Halpert and X. Wang, 1992. "Observed tropospheric biennial variability and its relationship to the Southern Oscillation," Journal of Climate, 5, 594-614.
- Sabbagh, M.E. and R.A. Bryson, 1962. "Aspects of the precipitation climatology of Canada investigated by the method of harmonic analysis," Annals of the Association of American Geographers, 52(4), 426-440.
- Salby, M.L. and D.R. Shea, 1991. "Correlations between solar activity and the atmosphere: an unphysical explanation," Journal of Geophysical Research, 96(D12), 22579-22595.
- Sangoyomi, T.B., U. Lall and H.D.I. Abarbanel, 1996. "Nonlinear dynamics of the Great Salt Lake: Dimension estimation," Water Resources Research, 32(1), 149-159.
- Schultz, D.M.; W.E. Bracken, L.F. Bosart, G.J. Hakim, M.A. Bedrick, M.J. Dickenson and K.R. Tyle, 1997. "The 1993 superstorm cold surge: Frontal structure, gap flow, and extratropical-tropical interaction," Monthly Weather Review, 125, 5-39.
- Schultz, D.M., W.E. Bracken and L.F. Bosart, 1998. "Planetary- and synoptic-scale signatures associated with Central American cold surges," Monthly Weather Review, 126, 5-27..

- Schwartz, G., 1978. "Estimating the dimension of a model," The Annals of Statistics, 6, 461-464.
- Sharp, G.D., 1992. "Fishery catch records, El Niño/Southern Oscillation, and longer-term climate change as inferred from fish remains in marine sediments," in El Niño: Historical and Paleoclimatic Aspects of the Southern Oscillation (H.F. Diaz and V. Markgraf, eds.). Cambridge, England: Cambridge University Press.
- Shaw, G. and D. Wheeler, 1985. Statistical Techniques in Geographical Analysis. New York, NY: John Wiley and Sons.
- Simpson, J.J., 1984. "A simple model of the 1982-1983 California 'El Niño'," Geophysical Research Letters, 11, 137-240.
- Smith, J.A., 1987. "Statistical modeling of daily rainfall occurrences," Water Resources Research, 23, 885-893.
- Statsoft, 1995. Statistica, Volume III. Tulsa, OK: Statsoft, Inc.
- Stedinger, J.R., 1980. "Fitting log normal distributions to hydrologic data," Water Resources Research, 16, 481-490.
- Steiner, E.J. and S.J.S. Khalsa, 1987. "Sea surface temperature, low-level moisture, and convection in the tropical Pacific, 1982-1985," Journal of Geophysical Research, 92, 14217-14224.
- Stern, R.D. and R. Coe, 1983. "A model fitting analysis of daily rainfall data," Journal of the Royal Statistical Society, 147, 1-34.
- Sumner, G., 1996. "Daily precipitation patterns over Wales: Towards a detailed precipitation climatology," Transactions of the Institute of British Geography, NS 21, 157-176.
- Tinsley, B.A., 1997. "Do effects of global atmospheric electricity on clouds cause climate changes?" EOS, 78(33), 341-ff.
- Trenberth, K.E., 1990. "Recent observed interdecadal climate changes in the northern hemisphere," Bulletin of the American Meteorological Society, 71(7), 988-993.
- Umatani, S. and T. Yamagata, 1991. "Response of the eastern tropical Pacific to meridional migration of the ITCZ: The generation of the Costa Rica Dome," Journal of Physical Oceanography, 21, 346-363.
- Van Loon, H. and R.A. Madden, 1981. "The Southern Oscillation. Part 1: Global associations with pressure and temperature in northern winter," Monthly Weather Review, 109, 1150-62.

Van Loon, H. and K. Labitzke, 1990. "Association between the 11-year solar cycle and the atmosphere. Part IV: The stratosphere, not grouped by phase of the QBO," Journal of Climate, 3, 827-837.

Vargas, A.B. and V.F.S. Trejos, 1994. "Changes in the general circulation and its influence on precipitation trends in Central America: Costa Rica," Ambio, 23, 87-90.

Vautard, R., P. Yiou and M. Ghil, 1992. "Singular-spectrum analysis: A toolkit for short, noisy chaotic signals," Physica D, 58, 95-126.

Vogel, R.M. and N.M. Fennessey, 1993. "L moment diagrams should replace product moment diagrams," Water Resources Research, 29, 1745-1752.

Walker, G.T. and E.W. Bliss, 1932. "World weather V," Memoirs of the Royal Meteorological Society, 4, 53-84.

Wallis, J.R., N.C. Matalas and J.R. Slack, 1974. "Just a moment!" Water Resources Research, 10, 211-219.

Waylen, P.R., C.N. Caviedes and M.E. Quesada, 1996. "Interannual variability of monthly precipitation in Costa Rica," Journal of Climate, 9, 2606-2613.

Waylen, P.R., M.E. Quesada and C.N. Caviedes, 1994. "The effects of El Niño-Southern Oscillation in San José, Costa Rica," International Journal of Climatology, 14, 559-568.

West, R.C. and J.P. Augelli, 1989. Middle America: Its Lands and Peoples, 3rd Edition. Englewood Cliffs, New Jersey: Prentice Hall, 494 pp.

Wilby, R.L., 1994. "Stochastic weather type simulation for regional climate change impact assessment," Water Resources Research, 30, 3395-3403.

Wilks, D.S., 1993. "Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series," Water Resources Research, 29(10), 3543-3549.

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences. New York, New York: Academic Press, 467 pp.

Willeke, G.E., 1980. "Myths and uses of hydrometeorology in forecasting," in Improved Hydrologic Forecasting - Why and How, Proceedings of Engineering Foundation Conference, March 25-30, 1979. Pacific Grove, CA: American Society of Civil Engineers, 117-124.

Woolhiser, D.A., 1992. "Modeling daily precipitation - progress and problems," in Statistics in the Environmental and Earth Sciences (A.T. Walden and P. Guttorp, eds.) New York, New York: Halsted Press.

Woolhiser, D.A., T.O. Keefer and K.T. Redmond, 1993. "Southern Oscillation effects on daily precipitation in the southwestern United States," Water Resources Research, 29, 1287-1295.

Woolhiser, D.A. and J. Roldan, 1986. "Seasonal and regional variability of parameters for stochastic daily precipitation models: South Dakota, U.S.A." Water Resources Research, 22, 965-978.

Wyrski, K., 1975. "El Niño -- The dynamic response of the equatorial Pacific to atmospheric forcing," Journal of Physical Oceanography, 5, 572-584.

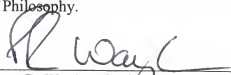
Wyrski, K., 1981. "An estimate of the equatorial upwelling in the Pacific," Journal of Physical Oceanography, 11, 1205-1214.

Wyrski, K., 1984. "The slope of sea level along the equator during the 1982-1983 El Niño," Journal of Geophysical Research, 89, 10419-10424.

BIOGRAPHICAL SKETCH

John Michael ("Mike") Harrison was born on March 18, 1960, the eldest son of Marshall Gene Harrison and Mary Ann Harrison. An "Air Force brat," he spent much of his youth in and around Air Force bases, eventually winding up in the Washington, D.C. area. Upon his father's retirement from the U.S. Air Force in 1973, the family relocated to Lubbock, Texas, where he lived until graduation from high school in 1978. He pursued an engineering degree at Rensselaer Polytechnic Institute in Troy, NY, attaining a Bachelor of Science degree in Computer and Systems Engineering in 1982. Following for awhile in his father's footsteps, he became an officer in the U.S. Air Force upon graduation, and was stationed at Los Angeles AFS, CA, where he worked as a space systems engineer on the Navstar Global Positioning System. He left the Air Force in 1988 to pursue graduate studies. He enrolled at the University of Georgia in 1990, where he worked with Dr. Chor Pang Lo, earning a Master of Arts degree in Geography in 1992. He entered the Geography program at the University of Florida in January 1994, where he has worked with Dr. Peter Waylen studying the characteristics of daily rainfall in Costa Rica. Upon graduation with his Ph.D., he will begin a tenure-track position at the University of Southern Mississippi in the Department of Geography.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Peter R. Waylen, Chairman
Professor of Geography

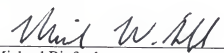
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


César N. Caviedes
Professor of Geography

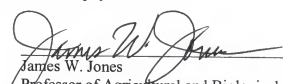
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Joann Mossa
Associate Professor of Geography

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Michael Binford
Associate Professor of Geography

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


James W. Jones
Professor of Agricultural and Biological Engineering

This dissertation was submitted to the Graduate Faculty of the Department of Geography in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1998

Dean, Graduate School